

Título: MANUAL DE METODOLOGIA E USO DO PREVECARGADESSEM 1.26.8	Nº de Páginas: 38 Nº de Anexos: 0
---	--

Dados do Cliente:

OPERADOR NACIONAL DO SISTEMA ELÉTRICO - ONS
Rua Júlio do Carmo, 251 - Cidade Nova
Rio de Janeiro - RJ, CEP: 20211-160

Responsável:

Maria Helena Teles de Azevedo
E-mail: mhelena@ons.org.br
Telefone: 21- 3444-9826

Departamento: Departamento de Otimização Energética e Meio Ambiente - DEA

Centro de Lucro: C230000101

Centro de Custo: C203000015

Resumo:

O presente relatório descreve a metodologia de previsão de carga implementada no Prev cargaDESSEM 1.26.8, apresenta os dados necessários para o cálculo da previsão e fornece informações para a execução do modelo em ambiente R.

Autores:

José Francisco Moreira Pessanha - Cepel
Victor Andrade de Almeida - Cepel

Palavras-Chave:

Previsão de carga
Machine Learning
DESSEM

Classificação de Acesso: Setorial

Aprovação e data de emissão

André Luiz Diniz Souto Lima
Chefe do Departamento DEA

Tel.: 21-2598-6046

E-mail: diniz@cepel.br

Aprovação e data de emissão

Maurício Barreto Lisboa
Diretor de Pesquisa, Desenvolvimento e Inovação

Sumário

SUMÁRIO EXECUTIVO	4
1. INTRODUÇÃO	5
2. MÉTODOS DE PREVISÃO USADOS NO PREVCARGADESSEM	7
2.1 Introdução	7
2.2 Redes Neurais Artificiais	7
2.3 Máquinas de vetor de suporte para regressão	13
2.4 Regressão quantílica	16
2.5 Previsão combinada	16
3. MODELAGEM DA PREVISÃO DE CARGA NO PREVCARGADESSEM	16
3.1 Introdução	16
3.2 Metodologia de tratamento de dados	17
3.2.1 Tratamento dos dados de temperatura	17
3.2.2 Tratamento dos dados de carga	18
3.3 Modelos de previsão de carga	20
3.3.1 Modelo de previsão para dias normais	20
3.3.2 Modelo de previsão para feriados e dias especiais	22
3.3.3 Avaliação a posteriori	25
4. PROGRAMA PREVCARGADESSEM	26
4.1 Instalação	26
4.2 Arquivos de dados	27
4.2.1 Histórico da carga global em base horária	27
4.2.2 Pesos das previsões no cálculo da previsão combinada	28
4.2.3 Lista de feriados e dias especiais	28
4.2.4 Horário de verão	29
4.2.5 Dados do horizonte de previsão	29
4.2.6 Patamares de carga	30
4.2.7 Separador decimal	30
4.2.8 Histórico das temperaturas horárias	30
4.2.9 Histórico das temperaturas horárias	31
4.2.10 Ponderação dos modelos de previsão	31
5. EXECUÇÃO EM AMBIENTE R	32

5.1 Execução no modo individual	32
5.2 Execução no modo combinado	35
REFERÊNCIAS BIBLIOGRÁFICAS	36

Sumário Executivo

O programa PrevcargaDESSEM, desenvolvido pelo Cepel, fornece previsões de carga até 192 horas à frente. O programa está disponível aos agentes e as previsões geradas são diariamente divulgadas no Portal Síntegra do ONS. Em 2020, a versão 1.26.8 do programa foi apresentada da 4ª FT PrevcargaDESSEM, na qual a formulação matemática dos modelos de previsão, baseadas em Support Vectors Machine (SVM) e Redes Neurais Artificiais, foi modificada com a inclusão de um parâmetro de suavização que leva em conta a idade dos padrões de treinamento. A motivação para a modificação proposta foi a alteração do perfil da carga decorrente da pandemia do novo coronavírus COVID 19. O presente relatório apresenta a nova formulação matemática dos modelos de previsão disponibilizados no PrevcargaDESSEM 1.26.8 e serve como manual de utilização do programa.

1. Introdução

As previsões de carga até 8 (oito) dias à frente constituem informações fundamentais para o Modelo de Despacho Hidrotérmico de Curto Prazo (DESSEM), modelo de referência para a definição do CMO e preços horários na comercialização de energia. As previsões devem ser expressas em MWmédio e fornecidas em intervalos semi-horários para os dois primeiros dias do horizonte de previsão e agregadas em patamares de carga leve, média e pesada (MWmédio por patamar) para os demais dias do horizonte de previsão.

A dinâmica e cronologia da previsão ocorre da seguinte forma: o horizonte de previsão é dinâmico e sempre se inicia com a produção de uma Função de Custo Futuro (FCF) pelo modelo DECOMP para o PMO ou suas revisões. Assim, toda quinta-feira é produzida uma nova FCF para a semana operativa que começa no sábado e termina na sexta-feira seguinte. Trata-se, portanto, de uma janela de tempo que se fecha até chegar em uma quinta-feira, onde uma nova FCF será produzida.

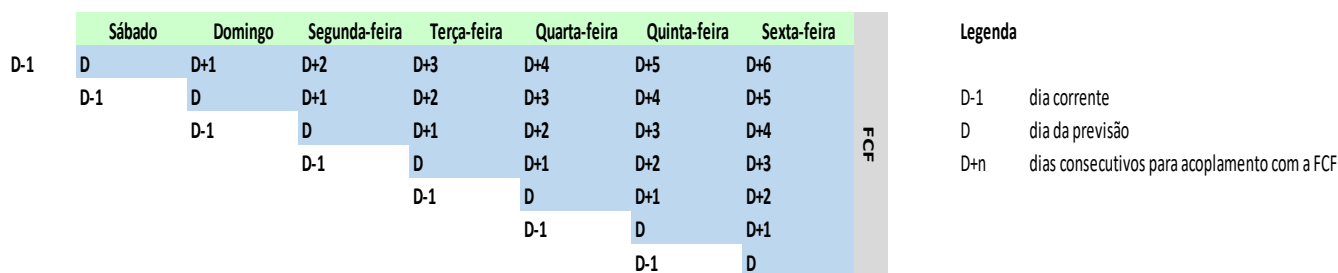


Figura 1. Dinâmica e Cronologia da Execução do PrecvargaDESSEM.

Seguindo a nomenclatura da Figura 1, para os dias "D" e "D+1" as previsões são disponibilizadas em intervalos semi-horários para cada subsistema, e para os demais dias até o final da semana operativa as previsões da carga são apresentadas em patamares (pesada, média e leve), conforme ilustrado pela curva em azul na Figura 2. Ainda na Figura 2, a linha em preto representa a previsão em base semi-horária. Inicialmente o PrecvargaDESSEM calcula as previsões em base horária e na sequência, por meio de interpolações por *splines* cúbicas (DUBINSKY & ELPERIN, 1997), gera as previsões semi-horárias para todo o horizonte de previsão. Por fim, as previsões semi-horárias são agrupadas por patamar de carga (PESSANHA et al, 2018).

De forma a automatizar o processamento da previsão de carga pelo ONS, o CEPEL desenvolveu o programa computacional PrecvargaDESSEM, cujas funções básicas incluem a leitura de dados (carga, temperatura, datas de início e fim do horário de verão, lista de feriados/dias especiais e intervalos horários dos patamares de carga), o cálculo das previsões da carga e a exportação das previsões para o modelo DESSEM.

Na Figura 3 tem-se uma ilustração do fluxo de processamento dos dados e obtenção das previsões pelo PrecvargaDESSEM. Conforme ilustrado na Figura 3, o modelo de previsão é alimentado com os dados horários de carga, além de dados horários de temperatura, datas de início e fim do horário de verão e a lista de feriados e dias especiais na Tabela 1. Na sequência por meio de modelos baseados em algoritmos de Máquina de Vetor de Suporte - *Support Vector Machine* (HAMEL, 2009) e Redes Neurais Artificiais (HAYKIN, 2009) são geradas previsões horárias da carga global. As previsões horárias passam por um processo de interpolação por *splines* cúbicas (DUBINSKY & ELPERIN, 1997) com a finalidade de obter previsões em base semi-horária e, por fim, as previsões são agrupadas nos períodos correspondentes aos patamares de carga indicados na Tabela 2 (PESSANHA et al, 2018).

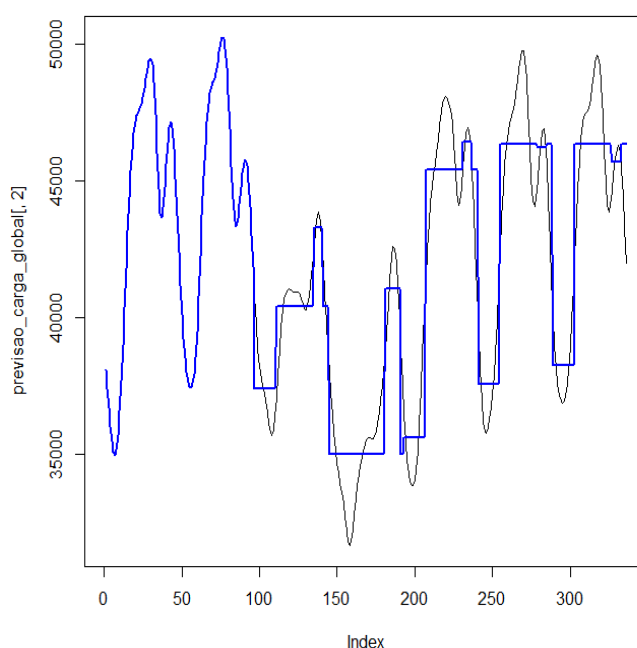


Figura 2. Previsão de carga com resolução semi-horária e por patamar.

Embora não exista um consenso acerca da superioridade de um determinado método de previsão de carga em particular, ressalta-se a grande quantidade de publicações que ilustram a aplicação de métodos de aprendizagem de máquina na previsão de carga horária, entre os quais se destacam as redes neurais artificiais (ZHANG et al, 1998; KODOGIANNIS & ANAGNOSTAKIS, 1999; HIPPERT et al, 2001; PARK et al, 2017), as máquinas de vetor de suporte (*support vector machine* – SVM) (CHEN et al, 2004; XU et al, 2005; NIU & WANG, 2007; WANG et al, 2007; HONG, 2009; BOZIC & STOJONAVIC, 2011; TURKAY et al, 2011; CEPERIC et al, 2013; BAGHEL et al, 2016) e a lógica *fuzzy* (MASTOROCOSTAS et al, 1999; LIAO & TSAO, 2004; PESSANHA & JUSTINO, 2014).

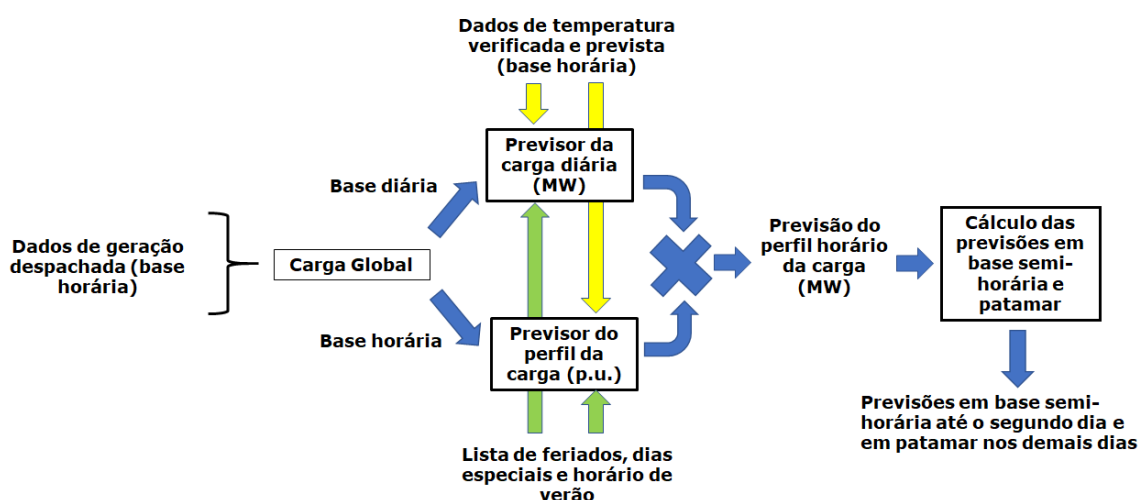


Figura 3. Estratégia de previsão de carga adotada no PrecvargaDESSEM.

Tabela 1. Codificação de feriados e dias especiais.

Tipo do dia	Código
Feriado Nacional perfil de sábado	1
Feriado Nacional perfil de domingo	2
Segunda-feira da semana do Carnaval	3
Terça-feira da semana do Carnaval	4
Quarta-feira de cinzas	5
Natal e 1º de janeiro	6
Dia após Natal ou após 1º de janeiro	7
Véspera de Natal ou véspera 1º de janeiro	8
Véspera de feriado	9
Dia após feriado	10
Dia especial e recesso	11
Blecaute, Jogos da Copa, etc.	12

Tabela 2. Definição dos patamares de carga (PESSANHA et al, 2018).

Patamar de carga	Inverno maio, junho, julho e agosto		Intermediário abril, setembro e outubro		Verão janeiro, fevereiro, março, novembro e dezembro	
	2ª a 6ª feira	sábado, domingo e feriado	2ª a 6ª feira	sábado, domingo e feriado	2ª a 6ª feira	sábado, domingo e feriado
Pesada	10h00-22h00		10h00-20h00		10h00-18h00	
Média	07h00-10h00 22h00-24h00	18h00-22h00	08h00-10h00 20h00-24h00	18h00-22h00	08h00-10h00 18h00-24h00	20h00-23h00
Leve	00h00-07h00	00h00-18h00 22h00-24h00	00h00-08h00	00h00-18h00 22h00-24h00	00h00-08h00	00h00-20h00 23h00-24h00

O presente relatório técnico tem por objetivo descrever a metodologia de previsão de carga adotada no PrevcargaDESSEM e fornecer instruções básicas para sua utilização e visualização dos resultados.

2. Métodos de previsão usados no PrevcargaDESSEM

2.1 Introdução

O PrevcargaDESSEM baseia-se em algoritmos estatísticos e de aprendizagem de máquinas, entre os quais destacam-se as Redes Neurais Artificiais - RNA e as Máquinas de Vetor de Suporte - SVM, mas também inclui em sua programação lógica fuzzy, modelo de regressão quantílica, análise de agrupamentos, estimador de Nadaraya-Watson, classificador Naive Bayes e modelo de programação quadrática. Nesta seção serão introduzidas algumas noções básicas dos métodos supracitados.

2.2 Redes Neurais Artificiais

O cérebro humano é composto por cerca de 100 bilhões de células nervosas, conhecidas por neurônios, que se conectam massivamente umas às outras através de ligações eletroquímicas, denominadas sinapses, formando uma grande rede de processamento. Cada neurônio recebe estímulos através dos dendritos, os processa em seu corpo celular e, dependendo do seu estado de ativação, gera e transmite um estímulo pelo seu axônio para que atinja outros neurônios. A estrutura e o funcionamento do neurônio biológico podem ser modelados pelo neurônio artificial ilustrado na Figura 4.

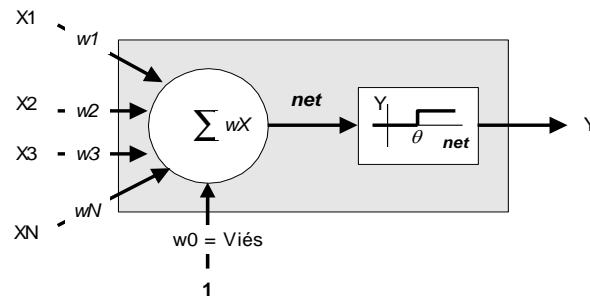


Figura 4. Neurônio artificial de McCulloch & Pitts (KOVÁCS, 1996).

No neurônio artificial os N terminais de entrada representam os dendritos e o único terminal de saída representa o axônio. A intensidade das sinapses é representada pelos pesos (w) nos terminais de entrada, cujos valores podem ser negativos ou positivos, definindo sinapses inibidoras ou estimuladoras respectivamente. De forma análoga ao que acontece no cérebro humano, onde as sinapses são reforçadas ou enfraquecidas, os pesos são ajustados durante a evolução do processo de aprendizagem da rede.

O comportamento do corpo celular é emulado por um mecanismo que faz a soma ponderada dos valores recebidos (cálculo do net). No modelo mais simples, esta soma ponderada é comparada com um limiar θ . Neste modelo, a função de ativação que relaciona a resposta do neurônio com a soma ponderada dos valores recebidos é uma função degrau: se $x_1 w_1 + x_2 w_2 + \dots + x_N w_N \geq \theta$ o neurônio é ativado e produz uma saída igual a 1, caso contrário o neurônio não dispara e a saída é zero. Enfim, o corpo celular é responsável pelo mapeamento dos sinais de entrada em um único sinal de saída. No lugar da função degrau pode-se usar funções contínuas, por exemplo, a função logística.

As RNA são sistemas paralelos distribuídos, compostos por unidades de processamento simples (neurônios) dispostas em uma ou mais camadas que são interligadas por muitas conexões geralmente unidirecionais e com pesos para ponderar a entrada recebida por cada neurônio. Através de uma analogia com o cérebro humano, pode-se afirmar que os pesos das conexões armazenam o conhecimento ou a memória da rede neural.

A organização de vários neurônios artificiais em uma estrutura e a forma de como eles se interligam define a arquitetura de uma RNA. A arquitetura mais usual é a rede *perceptron* de múltiplas camadas ou *Multilayer Perceptron* (MLP) com três camadas, conforme mostra a Figura 5.

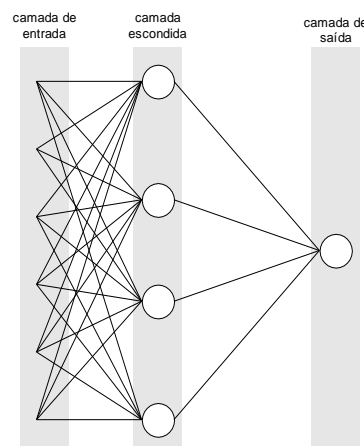


Figura 5. Rede MLP com três camadas.

A primeira camada da rede é a camada de entrada, sendo a única camada exposta aos sinais externos. Esta camada transmite os sinais recebidos para os neurônios da próxima camada, conhecida como camada escondida. Os neurônios das camadas escondidas extraem as características relevantes ou padrões dos sinais de entrada e transmitem o resultado para a camada de saída, a última camada da rede. A definição da arquitetura envolve a definição dos seguintes parâmetros: número de camadas escondidas, número de neurônios em cada camada e o tipo de conexão entre os neurônios.

A definição do número de camadas escondidas e do número de neurônios em cada camada é feita de forma empírica e depende da distribuição dos padrões de entrada, quantidade de ruído presente nos exemplos de treinamento e da complexidade da função a ser aprendida (HAYKIN, 2009).

O ajuste dos pesos sinápticos é realizado por meio de um processo iterativo denominado aprendizagem. O aprendizado pode acontecer, basicamente, de duas formas distintas: supervisionado e não supervisionado. O primeiro é comum em previsão de séries temporais, enquanto o segundo é utilizado para realizar análise de agrupamentos.

O aprendizado supervisionado consiste em apresentar padrões de entrada e saída desejada à RNA. Quando a saída gerada pela rede, a partir dos cálculos efetuados com o padrão de entrada e os pesos correntes, difere da saída desejada, os pesos são ajustados de forma a reduzir o erro. Essa dinâmica é repetida por todo conjunto de padrões (pares entradas e saídas) inúmeras vezes, até que a taxa de erro esteja dentro de uma faixa considerada satisfatória.

O método de treinamento mais usado no treinamento supervisionado é a retropropagação do erro ou *error backpropagation* (WERBOS, 1990), um algoritmo que utiliza pares (entrada, saída desejada) para através de um mecanismo de correção de erro ajustar os pesos da rede.

O ajuste dos pesos se dá pela execução de um processo de otimização realizado em duas fases: fase *forward* e fase *backward*, conforme ilustrado na Figura 6, onde os pesos são definidos de forma a minimizar a soma dos quadrados dos erros:

$$E = \frac{1}{2} \sum_p \sum_{j=1}^n \xi_j (d_j^p - y_j^p)^2 \quad (1)$$

em que p é o número de padrões de treinamento (padrões de entrada e saída), n é o número de neurônios da camada de saída, d_j é a saída desejada para o j -ésimo neurônio da camada de saída, y_j é o resultado gerado pelo j -ésimo neurônio da camada de saída e ξ_j é o peso atribuído ao j -ésimo padrão entrada/saída do conjunto de treinamento, segundo um decaimento exponencial, i.e., os padrões mais novos recebem pesos maiores e os padrões mais antigos recebem pesos menores.

Vale destacar que os pesos ξ em (1) dependem de uma constante α ($0 \leq \alpha \leq 1$) informada pelo usuário. Na análise de séries temporais, os N padrões de entrada/saída são ordenados cronologicamente, assim o peso atribuído ao i -ésimo padrão de entrada é $\xi_i = (1 - \alpha)^{N-i+1} / \sum_{j=1}^N (1 - \alpha)^{N-j+1}$. Para $\alpha=0$, os pesos ξ atribuídos aos padrões entrada/saída são iguais a 1, i.e., as idades dos padrões entrada/saída não interferem no treinamento da rede neural artificial.

Na fase *forward* a rede é apresentada à um padrão de entrada e gera uma saída. Na fase *backward* a diferença (erro) entre a saída desejada e a saída gerada pela rede é utilizada para ajustar os pesos das sinapses. Para minimizar a soma dos quadrados dos erros o algoritmo de aprendizagem se baseia no método do gradiente descendente, por isso, a fim de que esse método seja utilizado a função de ativação precisa ser contínua, diferenciável e não decrescente, por exemplo, a função logística.

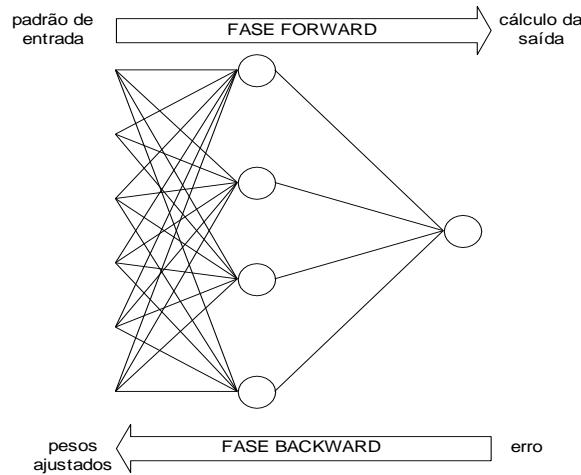


Figura 6. Fases *forward* e *backward* do algoritmo de retropropagação do erro.

O algoritmo *backpropagation* pode ser resumido nas seguintes operações (HAYKIN, 2009):

1º Passo - Inicialize aleatoriamente os pesos da rede e faça o contador de épocas (t) igual a zero.

2º Passo - Apresente uma época de exemplos de treinamento à rede. Uma época indica o número de vezes que os padrões de entrada (x) e o respectivo padrão de saída (y^d) (padrões do conjunto de treinamento) são apresentados à rede. Para cada exemplo de treinamento realize os passos 3 e 4 a seguir.

3º Passo – Fase forward: Propague o padrão de entrada (X_p) para frente, camada por camada, até chegar na camada de saída. Para cada neurônio calcule a combinação dos sinais recebidos da camada anterior:

$$net_j^I(t) = \sum_{i=0}^m w_{j,i}^I(t) \cdot y_i^{I-1}(t) \quad (2)$$

em que $y_i^{I-1}(t)$ é a saída produzida pelo neurônio i da camada anterior $I-1$ e $w_{j,i}^I(t)$ é o peso sináptico da conexão do neurônio j na camada I com o neurônio i da camada anterior $I-1$.

Para $i=0$ tem-se o viés (bias) aplicado ao neurônio j na camada I : $y_0^{I-1}(t)=1$ e $w_{j,0}^I(t)=b_j^I(t)$.

Se o neurônio j está na primeira camada oculta, i.e., $I=1$ $y_i^0(t)=x_j(t)$.

Denotando por f_j a função de ativação do neurônio j , o sinal produzido por ele é:

$$y_i^I(t) = f_j(net_j(t)) \quad (3)$$

No final calcule o erro, ou seja, a diferença entre a resposta desejada e a resposta fornecida pela rede:

$$e_j(t) = y_j^d - y_j \quad (4)$$

4º Passo – Fase backward: Propague o erro calculado no passo 3 para trás, começando na camada de saída e terminando na camada de entrada. Neste processo ajuste os pesos da seguinte forma:

$$w_{j,i}^I(t+1) = w_{j,i}^I(t) + \eta \delta_j^I(t) y_i^{I-1}(t) \quad (5)$$

em que η é a taxa de aprendizagem, $y_i^{I-1}(t)$ é a resposta do neurônio i situado na camada anterior $I-1$ e $\delta_j^I(t)$ é o gradiente local do neurônio j da camada I , definido de acordo com a localização do neurônio na rede.

Se a camada I onde está o neurônio é uma camada de saída tem-se:

$$\delta_j^I(t) = e_j(t) \cdot \frac{df_j(net_j(t))}{dnet_j} \quad (6)$$

Porém, se a camada I onde o neurônio estiver é uma camada escondida, o seu gradiente local é:

$$\delta_j^I(t) = \frac{df_j(net_j(t))}{dnet_j} \cdot \sum_{k=1}^m \delta_k^{I+1}(t) \cdot w_{kj}^{I+1}(t) \quad (7)$$

em que m é número de neurônios da camada $I+1$, $\delta_k^{I+1}(t)$ é o gradiente local do neurônio k situado na camada $I+1$ e $w_{kj}^{I+1}(t)$ é o peso sináptico da conexão entre o neurônio j na camada I e o neurônio k na camada $I+1$

5º Passo – Após terminar uma época de exemplos faça $t=t+1$ e itere para frente e para trás os passos 3 e 4. Pare apenas quando o critério de parada¹ for satisfeito.

A construção de um modelo baseado em rede neural artificial envolve a busca da melhor configuração para a rede, i.e., a definição do número adequado de camadas escondidas e a definição da quantidade ideal de processadores nestas camadas. A inexistência de regras que definam claramente a configuração adequada faz desta busca um processo empírico que demanda um longo tempo de processamento, em que diferentes configurações devem ser avaliadas. Entretanto, deve-se sempre empregar o princípio da parcimônia, e saber que uma rede com apenas uma camada oculta é capaz de aproximar qualquer tipo de função contínua (HAYKIN, 2009), embora em algumas situações específicas sejam usadas duas camadas.

Embora não existam leis que definam a configuração ótima, existem algumas propriedades que regem a convergência do algoritmo de treinamento, bem como a capacidade de generalização das RNA. O conhecimento destas propriedades orienta o processo de busca da configuração ótima e permite obter melhores resultados.

Em uma rede neural pequena há dificuldade de armazenar todos os padrões necessários e por isso a convergência do algoritmo de treinamento é mais lenta. Em uma rede pequena os processadores ficam sobrecarregados e lidam com muitas restrições na tentativa de encontrar uma representação ótima. Porém, deve-se ter o cuidado de não utilizar processadores demais, pois a rede pode memorizar os

¹ Usualmente o critério de parada fixa um determinado número de iterações ou estabelece uma tolerância para o erro.

padrões de treinamento, ao invés de extrair as características gerais que permitirão o reconhecimento de padrões não vistos durante o treinamento.

Com relação às redes com mais de uma camada escondida é importante observar que cada vez que o erro é propagado para a camada anterior, ele se torna menos útil e preciso. Apenas a camada imediatamente anterior à camada de saída tem uma noção precisa do erro, todas as camadas escondidas anteriores recebem uma estimativa do erro. Por esta razão a convergência dos pesos destas camadas é mais lenta.

No método *backpropagation*, o peso sináptico i do elemento processador j é atualizado de acordo com a seguinte regra entre as iterações t e $t+1$:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} \quad (8)$$

onde

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = -(\text{taxa de aprendizado}) \times \text{derivado do erro em relação a } w_{ij} \quad (9)$$

O *backpropagation* é inspirado no método do gradiente descendente: dado um ponto w no espaço dos pesos sinápticos, a taxa de aprendizagem η controla o tamanho do passo na direção oposta ao gradiente, a ser dado a partir deste ponto, visando minimizar o erro. Desta forma, um valor pequeno para η resulta em uma convergência lenta, enquanto um valor de maior magnitude para η acelera a convergência do treinamento, seja em redes com apenas uma camada escondida, ou em redes mais complexas, como a rede com duas camadas escondidas.

Entretanto, a taxa de aprendizagem não deve ser nem muito pequena e nem muito grande, pois com uma taxa pequena pode-se ficar preso em um mínimo local e a rede neural não conseguir aprender com a precisão especificada, por outro lado, valores grandes para a taxa de aprendizado tornam o treinamento instável, com fortes oscilações nos pesos de uma iteração para outra e o algoritmo de treinamento pode não convergir. Para acelerar o treinamento, porém evitando a instabilidade no ajuste dos pesos, adiciona-se o termo momento (α) no incremento dos pesos a cada iteração:

$$\Delta w_{ij}(t+1) = -\eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(t) \quad (10)$$

Nas regiões em que o gradiente aponta no mesmo sentido, a consideração do termo momento reforça o sinal dado pelo gradiente e acelera o treinamento, por sua vez, nas regiões em que o gradiente muda de direção, o termo momento suprime a oscilação nos pesos.

A previsão de valores futuros de uma série temporal, por meio de uma RNA (WERBOS, 1990, ZHANG et al, 1998), inicia-se com a montagem do conjunto de treinamento, que depende da definição do tamanho da janela de tempo para os valores passados das variáveis explicativas e do horizonte de previsão.

O padrão de entrada é formado pelos valores passados das variáveis explicativas e a saída desejada é o valor da observação da série temporal no horizonte de previsão. A Figura 7 ilustra como deve ser construído o conjunto de treinamento no caso da previsão um a passo à frente basear-se nos quatro últimos valores passados. A construção dos padrões de treinamento da rede consiste em mover as janelas de entrada e saída ao longo de toda série temporal. Cada par de janelas entrada/saída funciona

como um padrão de treinamento e deve ser apresentado repetidas vezes até que o algoritmo de aprendizado alcance a convergência.

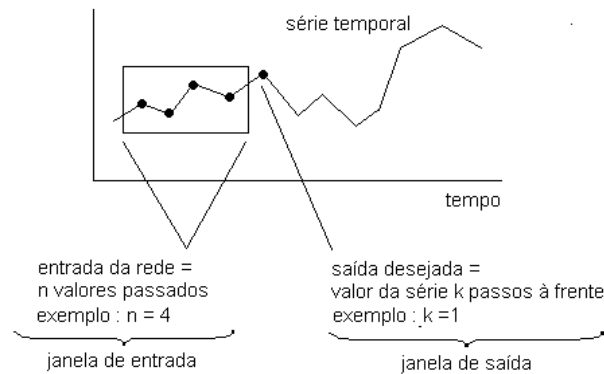


Figura 7. Montagem do conjunto de treinamento.

O padrão de entrada é formado pelos valores passados das variáveis explicativas e a saída desejada é o valor da observação da série temporal no horizonte de previsão. A Figura 7 ilustra como deve ser construído o conjunto de treinamento no caso da previsão um a passo à frente basear-se nos quatro últimos valores passados. A construção dos padrões de treinamento da rede consiste em mover as janelas de entrada e saída ao longo de toda série temporal. Cada par de janelas entrada/saída funciona como um padrão de treinamento e deve ser apresentado repetidas vezes até que o algoritmo de aprendizado alcance a convergência.

2.3 Máquinas de vetor de suporte para regressão

Proposta por Vapnik (1992), o algoritmo denominado Máquinas de Vetor de Suporte - SVM pode ser usado tanto em problemas de classificação de padrões quanto em problemas de regressão (SMOLA & SCHOLKOPF, 2004), neste último caso denomina-se Regressão por Vetores de Suporte (*Support Vector Regression – SVR*).

Considere uma série temporal com n observações de carga $y_i \forall i=1,n$ acompanhadas das respectivas variáveis explicativas representadas pelo vetor $X_i \forall i=1,n$. A previsão de carga pode ser efetuada pelo ajuste do seguinte modelo de regressão:

$$f(X) = w^T X + b \quad (11)$$

Conforme ilustrado na Figura 8a, a equação de regressão em (11) pode ser tratada como um hiperplano no centro de um hipertubo (*hypertube*) de largura 2ε com $\varepsilon > 0$ que envolve todas as observações.



(a) Hipertubo e hiperplano de regressão

(b) Variáveis de folga ξ

Figura 8. Hipertubo, hiperplano de regressão e variáveis de folga.

A melhor forma de alinhar o hipertubo aos dados consiste em posicioná-lo de maneira que as observações fiquem próximas da sua parede. Apesar do alinhamento ótimo do hipertubo aos dados não ser efetuado pelo método dos mínimos quadrados ordinários, o seu posicionamento é determinado pela solução do seguinte problema de programação quadrática:

$$\begin{aligned} \underset{w, b}{Min} \quad & \frac{1}{2} w^T w \\ & y_i - (w^T X_i + b) \leq \varepsilon \\ & (w^T X_i + b) - y_i \leq \varepsilon \end{aligned} \quad (12)$$

Contudo, não se pode garantir que todos os dados estejam dispostos dentro de um hipertubo de largura 2ε . Para acomodar estas situações o modelo em (12) deve ser modificado com a adição de variáveis de folga ξ indicadas no modelo em (13). Geometricamente, as variáveis de folga representam os desvios entre as paredes do hipertubo e as observações exteriores ao hipertubo, conforme indicado na Fig. 8b.

$$\begin{aligned} \underset{w, b, \xi^-, \xi^+}{Min} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^n \eta_i (\xi_i^+ + \xi_i^-) \\ & y_i - (w^T X_i + b) \leq \xi_i^+ + \varepsilon \\ & (w^T X_i + b) - y_i \leq \xi_i^- + \varepsilon \\ & \xi_i^-, \xi_i^+ \geq 0 \end{aligned} \quad (13)$$

No problema de otimização em (13) a constante C é positiva e permite controlar o *trade-off* entre a maximização da margem e a minimização das variáveis de folga, um mecanismo de prevenção ao sobreajuste (*overfitting*). Adicionalmente, um peso η_i é atribuído às variáveis de folga do i -ésimo padrão entrada/saída do conjunto de treinamento, segundo um decaimento exponencial, i.e., os padrões mais novos recebem pesos maiores e os padrões mais antigos recebem pesos menores. Vale destacar que os pesos η em (13) dependem de uma constante α ($0 \leq \alpha \leq 1$) informada pelo usuário. Na análise de séries temporais, os n padrões de entrada/saída são ordenados cronologicamente, assim o peso atribuído ao i -ésimo padrão de entrada é $\eta_i = (1 - \alpha)^{n-i+1} / \sum_{j=1}^n (1 - \alpha)^{n-j+1}$. Para $\eta=0$, os pesos atribuídos aos padrões entrada/saída são iguais a 1, i.e., as idades dos padrões entrada/saída não interferem no ajuste do SVM. A solução do problema de otimização em (13) na sua forma dual em (14) é mais simples de ser alcançada.

$$\begin{aligned} \underset{\lambda_i^+, \lambda_i^-}{Max} \quad & -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\lambda_i^+ - \lambda_i^-) (\lambda_j^+ - \lambda_j^-) X_i^T X_j + \sum_{i=1}^n y_i (\lambda_i^+ - \lambda_i^-) - \varepsilon \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-) \\ \text{s. a.} \quad & \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-) = 0 \\ & 0 \leq \lambda_i^+, \lambda_i^- \leq C \eta_i \end{aligned} \quad (14)$$

em que λ denota os multiplicadores de Lagrange associados com as restrições do problema de otimização em (13).

As relações entre as soluções dos problemas primal em (13) e dual em (14) são estabelecidas pelas equações (15) e (16) a seguir:

$$w = \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-) X_i \quad (15)$$

$$b = \frac{1}{n} \sum_{i=1}^n w^T X_i - y_i \quad (16)$$

Substituindo a equação (15) em (11) tem-se a seguinte equação de previsão para uma entrada X :

$$f(X) = \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-) X_i^T X + b \quad (17)$$

Uma vantagem do algoritmo SVR é a facilidade com que se pode estender o modelo de regressão linear no modelo de regressão não linear. A transformação da equação de previsão em (17) e do problema de programação matemática em (14) para o caso não linear consiste em substituir o produto $X_i^T X_j$ por uma função núcleo (*kernel trick*) que mapeia o espaço de entrada em um espaço de maior dimensão:

$$\Phi(X_i, X_j) = \phi(X_i)\phi(X_j) \quad (18)$$

O modelo de regressão resultante é ilustrado a seguir em (19). Na sequência, em (20) tem-se a equação de previsão para o caso não linear:

$$\begin{aligned} \text{Max}_{\lambda_i^+, \lambda_i^-} \quad & -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\lambda_i^+ - \lambda_i^-) (\lambda_j^+ - \lambda_j^-) \Phi(X_i^T X_j) + \sum_{i=1}^n y_i (\lambda_i^+ - \lambda_i^-) - \varepsilon \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-) \\ \text{s. a.} \quad & \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-) = 0 \\ & 0 \leq \lambda_i^+, \lambda_i^- \leq C \eta_i \end{aligned} \quad (19)$$

$$f(X) = \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-) \Phi(X_i^T X) + b \quad (20)$$

Os resultados de uma Máquina de Vetores de Suporte para regressão são sensíveis à definição dos seus hiperparâmetros, um conjunto formado pela constante C , pela largura do tubo ε , pela função núcleo (*kernel*) e seus os parâmetros.

A definição dos valores ótimos para os hiperparâmetros, em especial a constante C e a largura ε consiste em aplicar a validação cruzada ou *k-fold cross validation* (YADAV & SHUKLA, 2016).

A *K-fold cross validation* consiste em dividir aleatoriamente o conjunto de dados original em um número k de subconjuntos de mesmo tamanho. Com isso, o treinamento é realizado k vezes, utilizando a cada rodada um dos subconjuntos para validação e todos os outros para o treinamento do modelo. Para cada subconjunto de treinamento são calculadas métricas de desempenho para avaliar o erro de generalização e, ao final, de todas as rodadas de treinamento, é calculado o valor médio dos erros. Uma vantagem da validação cruzada é que o treinamento é realizado em todo o conjunto de dados, esperando-se assim uma maior capacidade de generalização do modelo. A Figura 9 ilustra o processo de validação cruzada *K-fold* para $k=10$.

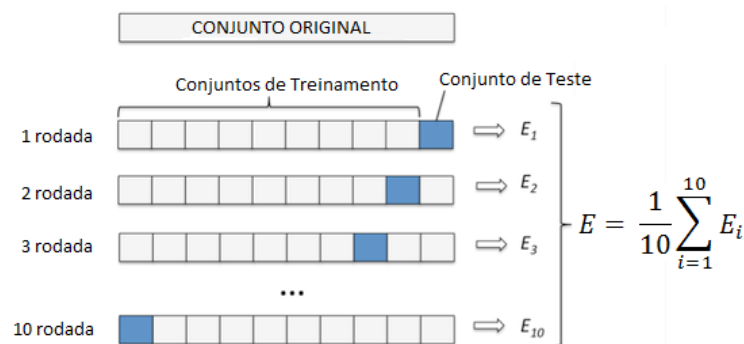


Figura 9. Exemplo de validação cruzada para $k=10$.

2.4 Regressão quantílica

Em um modelo de regressão linear, a equação de regressão estimada fornece uma estimativa do valor esperado da variável resposta y em função da variável explicativa X . Já em um modelo de regressão quantílica, a equação de regressão fornece uma estimativa do quantil τ (por exemplo, $\tau=0,5$ para a mediana) da variável resposta em função da variável explicativa. Assim, o conjunto de equações de regressão estimadas para diferentes valores de τ , permite caracterizar a distribuição de probabilidade da variável resposta condicionada aos valores das variáveis explicativas.

De forma distinta do estimador mínimos quadrados, na regressão quantílica para um quantil τ o estimador dos coeficientes de regressão busca minimizar a seguinte função objetivo em uma amostra com n pares $(y_i, X_i) \forall i=1, n$:

$$\underset{\beta}{Min} \sum_{i=1}^n \rho_{\tau}(y_i - X_i^T \beta) \quad (21)$$

em que $\rho_{\tau}(z) = z \cdot [\tau - I(z < 0)]$ e $I(z < 0)$ é uma função indicadora que assume valor unitário para $z < 0$, caso contrário, é igual a zero.

A função objetivo (21) pode ser inserida no seguinte problema de programação linear (PPL), cuja solução fornece as estimativas dos $K+1$ coeficientes de regressão:

$$\begin{aligned} \underset{u, v, \beta}{Min} \quad & \tau \sum_{i=1}^n u_i + (1 - \tau) \sum_{i=1}^n v_i \\ \text{s.a} \quad & y_i - X_i^T \beta = u_i - v_i \quad \forall i = 1, n \\ & u_i \geq 0 \quad \forall i = 1, n \\ & v_i \geq 0 \quad \forall i = 1, n \\ & \beta \in R^{K+1} \end{aligned} \quad (22)$$

Para $\tau = 0,5$ (mediana) o PPL em (22) corresponde ao estimador *Least Absolute Deviations* (LAD)

2.5 Previsão combinada

O PrevcargaDESSEM oferece 12 possibilidades para a previsão da demanda média diária, resultantes da combinação de 3 métodos de previsão (RNA, SVM com kernel linear e SVM com kernel radial) vezes 4 modelagens para a variável temperatura (sem temperatura, temperatura média, temperatura máxima e temperaturas mínima e máxima). O PrevcargaDESSEM permite que o usuário execute cada uma das 12 modelagens disponíveis e obtenha a previsão por cada método (previsão individual) ou faça uma execução em lote dos 12 modelos e ao final calcule a média ponderada das previsões (previsão combinada), com ponderações calculadas previamente pelo ONS. Bates e Granger (1969) mostram que a previsão combinada é melhor que as previsões individuais.

3. Modelagem da previsão de carga no PrevcargaDESSEM

3.1 Introdução

O modelo PrevcargaDESSEM produz previsões de carga até 192 horas à frente com resolução temporal semi-horária e por patamar. As previsões de carga são obtidas por meio de um processo organizado em cinco etapas:

Etapla 1: Leitura de dados.

Etapla 2: Tratamento de dados de carga e de temperatura

Etapa 3: Previsão do perfil horário da carga global em base horária.

Previsão de carga = $f(\text{carga passada, previsão de temperatura, feriados/dias especiais, período do ano})$ (23)

Etapa 3.1 Previsão da média diária da carga MW

Etapa 3.2 Previsão do perfil horário da carga em p.u.

Etapa 3.3 Previsão do perfil horário de carga em MW

Etapa 4: Cálculo das previsões em base semi-horária por meio de interpolação com *splines* cúbicas.**Etapa 5:** Cálculo das previsões de carga por patamar.

Ressalta-se que na Etapa 2 todos os dados de carga e temperatura são previamente tratados antes de serem utilizados pelas rotinas responsáveis pelo ajuste dos modelos de previsão. A metodologia de tratamento de dados é descrita na seção 3.2. Na sequência, o PrecvargaDESSEM calcula as previsões com base na metodologia descrita na seção 3.3. O PrecvargaDESSEM é um pacote desenvolvido para o ambiente "R" (R CORE TEAM, 2017). A interface do programa é realizada por meio de arquivos no formato "csv" (arquivo texto com colunas separadas por ",") ou por objetos "data.frame". O programa produz os seguintes resultados:

- Previsões de carga em base semi-horária/patamar para os subsistemas/áreas na Tabela 3.
- Arquivo com previsões de carga no formato do modelo DESSEM.

Tabela 3. Áreas e Subsistemas.

Subsistema SE/CO	Subsistema SUL	Subsistema NORDESTE	Subsistema NORTE
RJ – área RJ	RS – área RS	NEC – área Centro	AM – Amazonas
ES – área ES	SC – área SC	NEL – área Leste	PA – Pará
SP – área SP	PR – área PR	NES – área Sul	MA – Maranhão
MG – área MG	S – Subsistema SUL	NEO – área Oeste	TO – Tocantins
MT – área MT		NEN – área Norte	N – Subsistema Norte
DF – área DF		NESO – área Sudoeste	
GO – área GO		NE – Subsistema Nordeste	
MS – área MS			
AC – área AC			
RO – área RO			
SE – Subsistema SE/CO			

3.2 Metodologia de tratamento de dados**3.2.1 Tratamento dos dados de temperatura**

A temperatura é uma das principais variáveis explicativas da carga e neste sentido é fundamental que os dados de temperatura não apresentem valores discrepantes (*outliers*) e lacunas (*gaps*), pois a presença destes erros pode comprometer o ajuste do modelo de previsão. Inicialmente verifica-se a presença de lacunas de dados.

No caso de ausência de dados de temperatura ao longo de um dia inteiro (uma lacuna com duração de 24 horas), a lacuna é preenchida pelo perfil médio de temperatura resultante da média dos perfis diários imediatamente adjacentes ou pelo perfil imediatamente sem lacuna, no caso de um dos perfis adjacentes também ter lacuna. Em último caso, nos dias sem dados de temperatura, toma-se o perfil médio horário da temperatura ao longo do mesmo mês e ano do dia sem dados ou, no caso extremo, a média dos perfis horários do histórico com dados completos.

Na sequência, os eventuais *outliers* remanescentes são identificados e corrigidos com o auxílio do estimador de Nadaraya-Watson, um método de regressão não paramétrica (MARTINEZ & MARTINEZ, 2002). Dada uma série de temperatura, a respectiva versão suavizada é obtida por meio do seguinte estimador:

$$Y(t) = \frac{\sum_{i=1}^N K\left(\frac{t-t_i}{h}\right) y_i}{\sum_{i=1}^N K\left(\frac{t-t_i}{h}\right)} \quad (24)$$

em que K é uma função núcleo (gaussiana), N é o número de observações, t_i é o instante da temperatura verificada y_i e h é um hiperparâmetro denominado largura de banda.

Como resultado, o estimador de Nadaraya-Watson produz uma versão suavizada $Y(t)$ da série de temperatura. A partir do desvio padrão (DP) dos desvios entre as séries real e suavizada obtém-se os limites dos intervalos de confiança para a temperatura em cada hora:

$$\text{Limite superior: } LS(t) = Y(t) + 3,5 DP \quad (25)$$

$$\text{Limite inferior: } LI(t) = Y(t) - 3,5 DP \quad (26)$$

Os valores discrepantes situam-se fora dos intervalos de confiança definidos acima e são substituídos pelos valores suavizados correspondentes.

3.2.2 Tratamento dos dados de carga

Os dados de carga também são susceptíveis a lacunas e *outliers* e merecem passar por um tratamento prévio. O tratamento de dados de carga fornece uma versão tratada da série de carga com resolução horária.

3.2.2.1 Identificação dos perfis diários típicos

A metodologia de tratamento de dados de carga inicia-se com a identificação de lacunas de dados já durante a leitura de dados horários. As lacunas são substituídas por zero. Adicionalmente, os erros grosseiros (*outliers*) são identificados com o auxílio de um *boxplot* e os valores discrepantes são substituídos por zero.

Os perfis sem lacunas ou *outliers* são padronizados pelas respectivas médias diárias e posteriormente classificados pelo método de Ward (PESSANHA et al, 2015), um algoritmo de análise de agrupamentos, para a identificação automática de K perfis típicos de carga, em que cada perfil corresponde ao centroide de um *cluster* de perfis diários semelhantes. O algoritmo para agregação de N perfis diários de carga em K *clusters* (no PrecvargaDESSEM o valor de K foi fixado em 40) é descrito a seguir:

- 1) Inicie com N *clusters*, cada um contendo apenas um perfil diário de carga com resolução horária e construa a matriz de distâncias de ordem N entre os perfis.
- 2) Identifique o menor elemento da matriz de distâncias para encontrar o par de *clusters* mais similares.
- 3) Reúna os dois *clusters* identificados na etapa 2 em um único *cluster* e atualize a matriz de distâncias, retirando as linhas e colunas relativas aos dois *clusters* identificados em 2 e incluindo a linha e coluna com as distâncias entre os demais *clusters* e o novo *cluster* formado. Note que a ordem da matriz de distâncias diminui de uma unidade a cada vez que a etapa 3 é executada.

4) Repita os passos 2 e 3 até que restem apenas dois *clusters*. A cada iteração guarde a identificação dos *clusters* que foram fundidos e a distância entre eles.

Nas etapas 1 e 3 do algoritmo de Ward as distâncias entre os *clusters* são calculadas com base na seguinte fórmula:

$$d_{ij} = \frac{p_i p_j}{p_i + p_j} d^2(c_i, c_j) \quad (27)$$

em que p_i e p_j denotam as quantidades de objetos nos *clusters* i e j respectivamente e $d^2(c_i, c_j)$ representa o quadrado da distância euclidiana entre os centroides dos agrupamentos i e j .

3.2.2.2 Identificação do perfil diário para dias sem dados de carga

Ao final do método de Ward, cada perfil participante da análise de agrupamentos é rotulado com número entre 1 e K correspondente ao *cluster* em que foi alocado. Adicionalmente, para cada um destes perfis são conhecidos o dia da semana, o mês, o tipo de dia (normal, feriado ou dia especial) e o horário (normal ou verão). A partir destas informações podem ser calculadas as seguintes frequências (probabilidades condicionais):

$$P(\text{dia da semana } i \mid \text{cluster } j) = P(i \mid j) \quad \forall i=1,7 \text{ e } j=1,k \quad (28)$$

$$P(\text{mês } m \mid \text{cluster } j) = P(m \mid j) \quad \forall m=1,12 \text{ e } j=1,k \quad (29)$$

$$P(\text{tipo de dia } n \mid \text{cluster } j) = P(n \mid j) \quad \forall n=1,9 \text{ e } j=1,k \quad (30)$$

$$P(\text{horário } h \mid \text{cluster } j) = P(h \mid j) \quad \forall h=1 \text{ (verão)}, 2 \text{ (normal)} \text{ e } j=1,k \quad (31)$$

A probabilidade $P(\text{cluster } j) = P(j) \quad \forall j=1,k$ é a frequência relativa em cada *cluster*. As probabilidades acima podem ser combinadas pelo Teorema de Bayes com o intuito de identificar o perfil de carga típico (*cluster*) mais provável em um dia qualquer, conhecendo-se apenas a data (dia da semana e mês), o horário (normal ou verão) e o tipo do dia (normal, feriado ou dia especial). Admitindo que as características dos dias sejam condicionalmente independentes, o classificador *Naïve Bayes* (GRUS, 2016) atribui o perfil típico j ao dia da semana i , no mês m , com tipo n e horário (de verão e normal) h se a hipótese de *máxima a posteriori* for verificada para o *cluster* j , conforme descrito pela desigualdade a seguir:

$$P(i|j)*P(m|j)*P(n|j)*P(h|j)*P(j) > P(i|w)*P(m|w)*P(n|w)*P(h|w)*P(w) \quad \forall j \neq w=1,K \quad (32)$$

3.2.2.3 Estimativa da demanda média diária para os dias sem dados de carga

O perfil típico atribuído pelo classificador *Naïve Bayes* a um dia com perfil de carga incompleto está normalizado pela média, assim é preciso multiplicá-lo por uma estimativa da demanda média diária para obter o perfil horário em MW. A estimativa da demanda média diária para um dia D sem dados de carga é fornecida por uma Máquina de Vetor de Suporte em função das seguintes variáveis explicativas:

- valores passados da carga média diária nos dias $D-1$, $D-7$, $D-14$ e $D-21$
- temperatura máxima no dia D
- temperatura média no dia D
- temperatura mínima no dia D
- mês do dia D (11 variáveis indicadoras ou *dummy*, em dezembro todas as variáveis são nulas)
- dia da semana do dia D (6 variáveis indicadoras ou *dummy*, no sábado todas as variáveis são nulas)
- horário do dia D (0 para o horário normal e 1 para o horário de verão)
- tipo do dia (uma variável indicadora ou *dummy*) que assume valor 1 para os feriados e 0 nos demais dias.

O perfil estimado para um dia sem dados de carga (lacuna de 24 horas) é determinado pelo produto entre a previsão da demanda média diária fornecida pelo modelo SVM e o perfil (em p.u. da demanda média) atribuído pelo Naive Bayes.

3.2.2.4 Correção de valores discrepantes

Ao final, os eventuais *outliers* remanescentes, identificados pelo *boxplot* da série de carga, são substituídos pelos valores medianos na mesma hora e mesmo dia semana em uma janela de 744 horas. Na sequência aplica-se o estimador de Nadaraya-Watson (MARTINEZ & MARTINEZ, 2002) para suavizar a série de carga:

$$Y(t) = \frac{\sum_{i=1}^N K\left(\frac{t-t_i}{h}\right) y_i}{\sum_{i=1}^N K\left(\frac{t-t_i}{h}\right)} \quad (33)$$

em que K é uma função núcleo (gaussiana), N é o número de observações, t_i é o instante da carga verificada y_i e h é um hiperparâmetro denominado largura de banda.

Como resultado, o estimador de Nadaraya-Watson produz uma versão suavizada $Y(t)$ da série de carga. A partir do desvio padrão (DP) dos desvios entre as séries de carga e sua versão suavizada obtém-se os limites dos intervalos de confiança para a carga em cada hora:

$$\text{Limite superior: } LS(t) = Y(t) + 3,5 DP \quad (34)$$

$$\text{Limite inferior: } LI(t) = Y(t) - 3,5 DP \quad (35)$$

Os valores discrepantes situam-se fora dos intervalos de confiança definidos acima e são substituídos pelos valores suavizados correspondentes.

3.3 Modelos de previsão de carga

Para a previsão do perfil de carga em um dia D o PrevcargaDESSEM emprega modelos específicos para os dias normais e dias especiais/feriados.

3.3.1 Modelo de previsão para dias normais

Se o dia D da previsão é um dia normal, o PrevcargaDESSEM treina modelos SVM ou RNA para a previsão da demanda média diária em MW e na sequência treina 24 SVM, um para cada hora do dia, para previsão do respectivo perfil de carga em p.u. da demanda média diária. A previsão horária é obtida pelo produto entre a previsão da demanda média diária (MW) e a previsão do perfil horário (em p.u.).

Para a previsão da demanda média diária (em MW), em um dia D normal, precedido de dias normais em $D-1$, $D-2$ e $D-7$, o programa oferece três métodos: RNA (com dois neurônios em uma camada escondida), SVM com *kernel* linear e SVM com *kernel* radial. Nas três opções disponíveis as variáveis explicativas incluem as demandas médias dos dias $D-1$, $D-2$ e $D-7$ e a temperatura, sendo que para esta última variável são oferecidas quatro especificações possíveis: sem temperatura, média da temperatura diária, máxima temperatura diária e valores mínimo e máximo da temperatura diária. Portanto, o PrevcargaDESSEM oferece 12 modelagens possíveis para a previsão da demanda média diária em um dia D normal.

No caso de um dia D normal, precedido por dias normais em $D-1$, $D-2$ e $D-7$, em um mês M , o conjunto de treinamento envolve os padrões entrada/saída em todos os dias do histórico no mesmo mês M e nos meses imediatamente anterior e posterior ao mês M , portanto, uma janela Ω de 3 meses ao redor do mês M (Figura 10). Para o mês de janeiro a janela Ω inclui apenas os meses de janeiro e fevereiro. Assim, por exemplo, para a previsão de um dia D em junho, o conjunto de treinamento é

formado por padrões dos meses de maio, junho e julho coletados ao longo de todo o histórico. Naturalmente, os limites da janela devem respeitar o período do horário de verão ou horário normal dependendo do mês em que a previsão é realizada. Esta estratégia para seleção dos padrões entrada/saída é utilizada no treinamento dos modelos de previsão da demanda média diária e do perfil de carga em base horária para os dias normais, precedidos de dias normais. Ressalta-se que as observações relativas aos dias marcados com feriado tipo 12 não fazem parte do conjunto de treinamento.

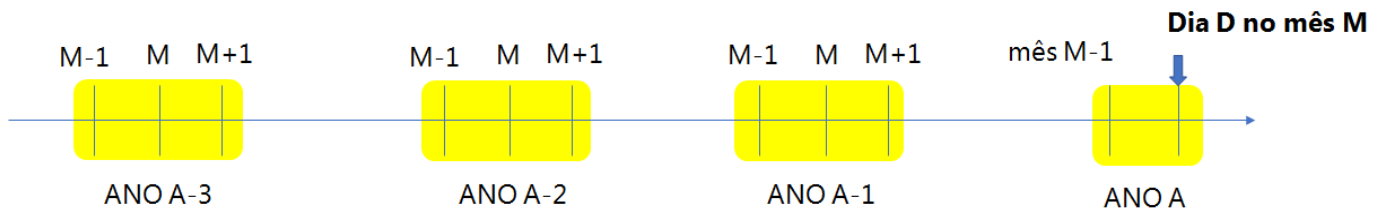


Figura 10. Janela trimestral para formação do conjunto de treinamento.

Para a previsão da demanda média diária, os padrões de saída correspondem aos valores da demanda média diária em todos os dias na janela Ω , mas apenas nos dias $d \in \Omega$ que ocorrem no mesmo dia da semana do dia D da previsão. Já os padrões de entrada incluem as demandas médias nos dias $d-1, d-2$ e $d-7$, dummy de feriados (nos dias $d-1, d-2$ e $d-7$), bem como a temperatura verificada no dia d (média, máxima ou mínima & máxima, conforme selecionado pelo usuário), caso a variável temperatura tenha sido incluída entre as variáveis explicativas.

Ainda na previsão da demanda média diária, caso a temperatura máxima prevista para o dia $D+1$ situe-se abaixo do quantil 1% ou acima do quantil 99% da distribuição das temperaturas máximas no conjunto de treinamento, a previsão da demanda média diária é fornecida por um modelo de regressão quantílica, ajustado para o quantil da temperatura máxima prevista com as mesmas variáveis consideradas nos modelos SVM e RNA.

Já a previsão do perfil horário de carga para um dia D normal, em p.u. da demanda média, conta com 24 SVM, um para cada hora $h \forall h=1,24$. Para a previsão da demanda horária na hora h , os padrões de saída correspondem aos valores da demanda da hora h em todos os dias na janela Ω , mas apenas nos dias $d \in \Omega$ que ocorrem no mesmo dia da semana do dia D da previsão. Já os padrões de entrada incluem as demandas horárias na hora h nos dias $d-7$ e $d-14$, a demanda média diária para o dia d , bem como a temperatura verificada na hora h do dia d , caso a variável temperatura tenha sido incluída entre as variáveis explicativas.

Caso o dia D seja normal, porém, os dias D-1, D-2 ou D-7 não sejam normais a previsão da demanda média diária para o dia D é realizada por RNA ou SVM com as seguintes variáveis explicativas:

- valores da demanda média diária nos dias D-1, D-2 e D-7
- tipo do feriado no dia D, 10 variáveis indicadoras (*dummy*), conforme indicado na Tabela 4. As observações relativas aos dias marcados com feriado tipo 12 não fazem parte do conjunto de treinamento. Nos dias normais todas as variáveis indicadoras são nulas.
- variável indicadora que assume valor 1 se dia D pertence ao período do horário de verão e 0 caso contrário.
- temperatura, segundo uma das quatro especificações possíveis: sem temperatura, média da temperatura diária no dia D, máxima temperatura diária no dia D e valores mínimo e máximo da temperatura diária no dia D.

Tabela 4. Variáveis indicadores dos tipos de feriados.

Dummy	Tipo do feriado									
	1	2	3/4	5	6	7	8	9	10	11
1	1									
2		1								
3			1							
4				1						
5					1					
6						1				
7							1			
8								1		
9									1	
10										1

Ressalta-se que neste último caso a previsão do perfil horário de carga em p.u. segue a mesma estratégia descrita acima para os dias normais precedidos por dias normais, i.e., envolve o ajuste de 24 SVM.

3.3.2 Modelo de previsão para feriados e dias especiais

Se o dia D da previsão for um feriado/dia especial o PrecvargaDESSEM faz uma previsão por similaridade, i.e., busca nos históricos de carga e temperatura um perfil de carga com características semelhantes às verificadas no momento da previsão.

Para ilustrar a ideia central da previsão por similaridade adotada no PrecvargaDESSEM considere a previsão do perfil de carga para o Natal (dia D da previsão), dado uma previsão do perfil de temperatura para o dia 25 de dezembro (temperatura no dia D) e que o perfil de carga do dia 24 de dezembro seja conhecido (carga no dia D-1).

Assim, busca-se no histórico todos os perfis de carga correspondentes ao mesmo feriado do dia D, neste caso o Natal. Cada perfil identificado forma o consequente de uma regra de associação do tipo "Se Antecedente Então Consequente".

Por sua vez, o antecedente de cada regra é formado pelo perfil de temperatura prevista para o dia D e o perfil de carga no dia D-1.

Para um histórico com três anos de dados há apenas três regras para o Natal, cujos antecedentes (temperatura em D e carga em D-1) e consequentes (carga em D) são apresentados graficamente na Figura 11.

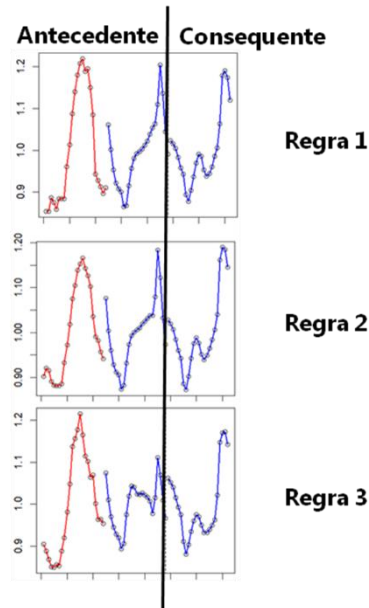


Figura 11. Regras de associação para o Natal.

Conforme ilustrado na Figura 12, para o dia D da previsão tem-se apenas os elementos que correspondem aos antecedentes das regras, i.e., a carga do dia anterior (D-1), em azul, e a previsão de temperatura para o dia D, em vermelho.

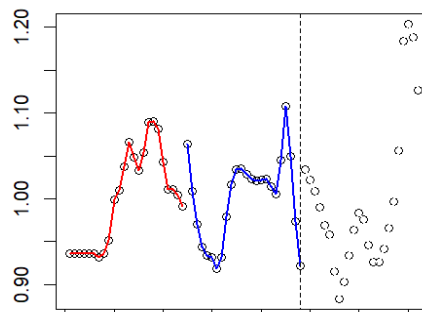


Figura 12. Antecedentes das regras para previsão do Natal.

A previsão obtida por similaridade é a média dos consequentes das três regras (perfis C1, C2 e C3) ilustrados na Figura 11, ponderados pelos respectivos graus de similaridade (w) entre o antecedente do dia D (perfil X) e os antecedentes das regras (perfis A1, A2 e A3).

$$\frac{w_1 C_1 + w_2 C_2 + w_3 C_3}{w_1 + w_2 + w_3} \quad (38)$$

O grau de similaridade entre o antecedente do dia D e o antecedente de uma regra (A1, A2 e A3) é calculado da mesma forma que o grau de pertinência no algoritmo *Fuzzy C-Means* (CANNON et al, 1986), um algoritmo para análise de agrupamentos:

$$w_i = \frac{1}{\left(\sum_{k=1}^3 \frac{\|X - A_i\|}{\|X - A_k\|} \right)^{\frac{2}{m-1}}} \quad \forall i = 1, 3 \quad (39)$$

Na equação (39) m é a constante de *fuzzyficação* e no PrevcargaDESSEM foi fixada em 1,25.

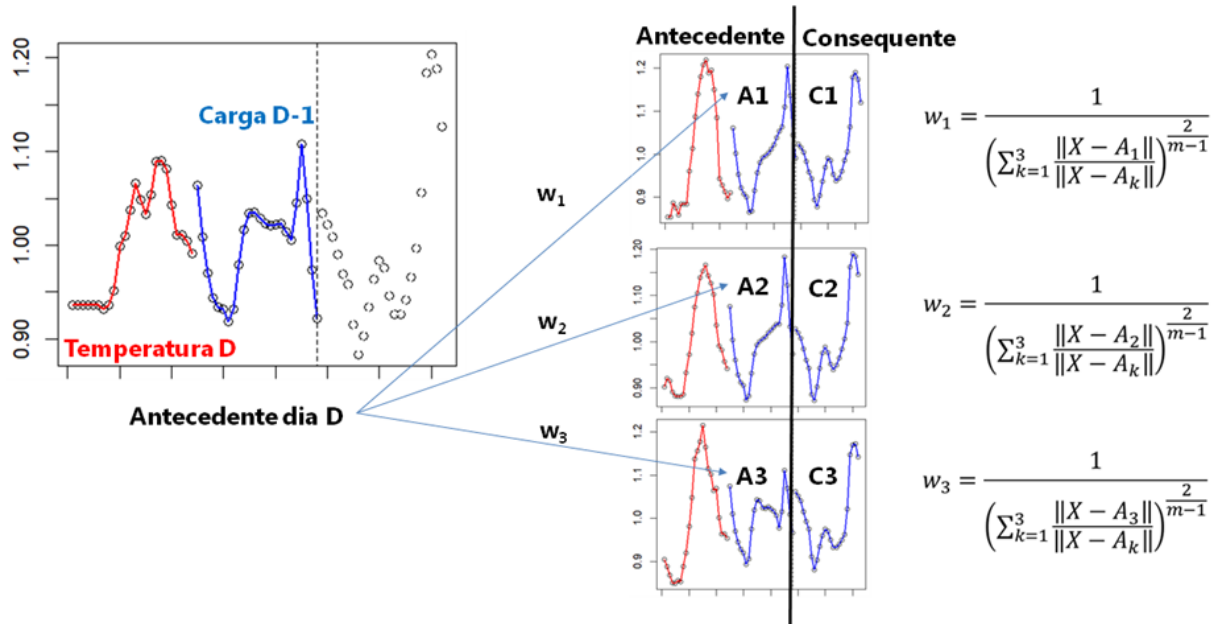


Figura 13. Antecedentes das regras para previsão do Natal.

Portanto a previsão do dia D é dada pela média ponderada em (38). Conforme ilustrado na Figura 13, para o dia D da previsão tem-se apenas os elementos que correspondem aos antecedentes das regras, i.e., a carga do dia anterior (D-1), em azul, e a previsão de temperatura para o dia D, em vermelho.

Vale destacar que no exemplo ilustrativo o antecedente de cada regra de associação é formado pelo perfil da temperatura prevista para o dia D e pelo perfil de carga para o dia D-1. A mesma especificação é usada na previsão dos feriados tipos 1, 2 e 6. Nos demais tipos de feriados e dias especiais o antecedente é formado pelo perfil da temperatura prevista para o dia D e pelo perfil de carga para o dia D-7. Exceto para os feriados tipo 6, 7 e 8 (relativos ao período do Natal e 1º de janeiro), para os demais feriados a formação do conjunto de regras de associação pode abranger feriados em uma janela de três meses centrada no mês do dia da previsão.

A ideia central do previsor por similaridade pode ser implementada de diferentes formas e no PrevcarregaDESSEM são avaliadas duas maneiras distintas e o próprio programa escolhe a melhor delas. Inicialmente, o PrevcarregaDESSEM aplica a ideia do previsor por similaridade nos perfis padronizados pela média e a previsão resultante é multiplicada pela previsão da demanda média oriunda de um modelo (RNA ou SVM) descrito na seção 3.3.2.1. Na sequência o PrevcarregaDESSEM aplica a mesma ideia do previsor por similaridade diretamente aos dados em MW, portanto, sem normalizá-los. Assim, ao final tem-se duas previsões e por se tratar de previsões para feriados, o programa escolhe a previsão que apresentar o perfil de carga em MW mais próximo do último sábado (no caso de feriado tipo 1) ou do último domingo (demais feriados).

3.3.2.1 Modelo para previsão da demanda média diária para feriados e dias especiais

A demanda média diária para feriados e dias especiais conta com duas modelagens distintas: uma modelagem simplificada para os tipos de feriados 1 e 2 e outra para os demais feriados. Para ambos os casos o usuário pode escolher uma RNA (com dois neurônios na camada escondida), um SVM com *kernel* linear ou um SVM com *kernel* radial.

Para os feriados tipo 1 e 2 o PrevcarregaDESSEM treina um modelo específico para a previsão da demanda média nos feriados com perfis de sábado e domingo. O conjunto de treinamento desse modelo é formado apenas por perfis de feriados tipo 1 e 2 presentes no histórico. Neste caso, a previsão da demanda média diária para o dia D é função das seguintes variáveis explicativas:

- valores da demanda média diária nos dias D-1, D-2 e D-7
- tipo do feriado no dia D, uma variável indicadora (*dummy*) que assume valor 1 para feriado com perfil de sábado e 0 para feriado com perfil de domingo. As observações relativas aos dias marcados com feriado tipo 12 não fazem parte do conjunto de treinamento.
- variável indicadora que assume valor 1 se dia D é feriado em dia útil e 0 caso contrário.
- variável indicadora que assume valor 1 se dia D pertence ao período do horário de verão e 0 caso contrário
- temperatura, segundo uma das quatro especificações possíveis: não considera temperatura, média da temperatura diária no dia D, máxima temperatura diária no dia D e valores mínimo e máximo da temperatura diária no dia D.

Já no caso dos demais tipos de feriados a previsão da demanda média é obtida por um outro modelo ajustado separadamente para cada dia da semana. Neste caso, a previsão da demanda média diária para o dia D é função das seguintes variáveis explicativas:

- valores da demanda média diária nos dias D-1, D-2 e D-7
- tipo do feriado no dia D, 10 variáveis indicadoras (*dummy*), conforme indicado na Tabela. As observações relativas aos dias marcados com feriado tipo 12 não fazem parte do conjunto de treinamento. Nos dias normais todas as variáveis indicadoras são nulas.
- variável indicadora que assume valor 1 se dia D pertence ao período do horário de verão e 0 caso contrário.
- temperatura, segundo uma das quatro especificações possíveis: não considera temperatura, média da temperatura diária no dia D, máxima temperatura diária no dia D e valores mínimo e máximo da temperatura diária no dia D.

Tabela 5. Variáveis indicadores dos tipos de feriados.

Dummy	Tipos de feriado									
	1	2	3/4	5	6	7	8	9	10	11
1	1									
2		1								
3			1							
4				1						
5					1					
6						1				
7							1			
8								1		
9									1	
10										1

3.3.3 Avaliação a posteriori

As previsões dos perfis de carga passam por uma avaliação a posteriori com o objetivo de avaliar a sua razoabilidade. A avaliação consiste na aplicação de um conjunto de heurísticas.

3.3.3.1 Avaliação a posteriori das previsões para feriados e dias especiais

As previsões dos perfis de carga para os feriados e dias especiais passam por uma avaliação a posteriori. Caso o dia D seja um feriado ou dia especial busca-se no histórico todos os dias com o mesmo tipo de feriado do dia D para formar um conjunto de padrões, os padrões selecionados podem pertencer a qualquer ano do histórico, mas devem ser do mesmo mês, ou meses imediatamente adjacentes ao mesmo mês do dia D de previsão. Na sequência ajusta-se um modelo de regressão linear simples em que a média da carga no dia do feriado é função da média da carga no dia imediatamente anterior ou na semana anterior. A partir da equação de regressão estimada pode-se obter um intervalo de previsão de 90% e uma nova estimativa para a demanda média diária no dia D da previsão. De forma resumida, se a demanda média prevista inicialmente para o dia D (previsão

calculada na seção 3.3.2) estiver fora dos limites do intervalo de confiança, uma nova previsão da demanda média para o dia D é definida pelo limite do intervalo de previsão que estiver mais próximo da previsão inicial (previsão calculada na seção 3.3.2), caso contrário a nova previsão da demanda média para o dia D é definida pela média entre a previsão inicial e a estimativa gerada pelo modelo de regressão.

Na sequência, com os mesmos dados selecionados no histórico e usados na regressão descrita acima calcula-se θ , i.e., a média da razão da demanda média diária entre dois dias consecutivos e o respectivo intervalo com 95% de confiança. Se a razão entre a previsão inicial da demanda média do dia D (previsão calculada na seção 3.3.2) pela demanda média do dia D-1 estiver fora dos limites do intervalo de confiança, a nova previsão da demanda média diária para o dia D é igual a θ vezes a demanda média do dia D-1.

3.3.3.2 Correção de degrau entre as previsões de dois dias consecutivos

Eventualmente a previsão para a primeira hora do dia D pode superar a previsão para a última hora do dia D-1, i.e., há um degrau indevido entre as previsões. A correção deste problema pode ser efetuada pelo modelo de programação quadrática em (40), cujo objetivo consiste em encontrar uma previsão x_h ($\forall h=1,24$) muito próxima do perfil de carga previsto p_h ($\forall h=1,24$) para o dia D, mas que não apresente um degrau em relação ao dia D-1.

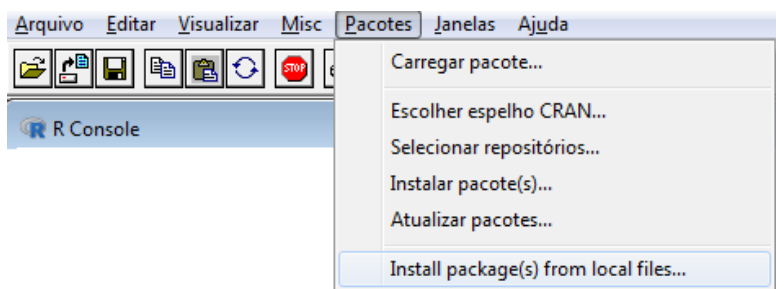
$$\begin{aligned}
 & \underset{x_h \forall h=1,24}{\text{Min}} \sum_{h=1}^{24} (x_h - p_h)^2 \\
 & \alpha_h \leq \frac{x_h}{x_{h-1}} \leq \beta_h \quad \forall h=1,24 \\
 & 0,98\bar{p} \leq \frac{1}{24} \sum_{h=1}^{24} x_h \leq 1,02\bar{p}
 \end{aligned} \tag{40}$$

Em (40) os limites das 24 restrições para a razão x_h/x_{h-1} ($\forall h=1,24$) são definidos com base na previsão inicial p_h ($\forall h=1,24$) e em x_0 , i.e., a carga do último intervalo horário do dia anterior (D-1). Estas restrições visam garantir que o perfil de carga resultante da otimização x_h ($\forall h=1,24$) tenha variações horárias próximas das encontradas na previsão inicial. Por fim, a última restrição visa garantir que a demanda média prevista da curva resultante da otimização seja próxima de \bar{p} , i.e., a demanda média inicialmente prevista para o dia D.

4. Programa PrecvargaDESSEM

4.1 Instalação

Para instalar o PrecvargaDESSEM copie o arquivo prevcargadessem.zip para algum diretório. Na sequência, acesse o R e use a opção do menu "Pacotes/Instalar pacote(s) a partir de arquivos zip locais..." (Figura 14) para instalar o pacote prevcargadessem disponibilizado no arquivo prevcargadessem.tar.gz. Adicionalmente, os pacotes e1071, WeightSVM, nnet, quantreg e quadprog devem ser previamente instalados.

Figura 14. Menu Pacotes e opção *Install packages(s) from local files*.

4.2 Arquivos de dados

Os arquivos com os dados de entrada para o PrevcargaDESSEM são disponibilizados no Portal Sintegre do ONS (<https://sintegre.ons.org.br>). Os dados estão disponíveis em arquivos zip, datados e disponíveis sob o título "decks para previsão de carga para o DESSEM" para cada subsistema. Em cada arquivo zip há 9 arquivos no formato csv, conforme ilustrado na Tabela 6, em que XX denota o acrônimo identificador da carga e AAAA-MM-DD a data do arquivo no formato ano/mês/dia.

Tabela 6. Arquivos de entrada do PrevcargaDESSEM.

Nome do arquivo	Conteúdo
XX_AAAA-MM-DD_CARGAHIST.CSV	Dados históricos da carga global em base horária
XX_AAAA-MM-DD_COMBINA.CSV	Pesos das previsões no cálculo da previsão combinada
XX_AAAA-MM-DD_FERIADOS.CSV	Datas de feriados nacionais e estaduais
XX_AAAA-MM-DD_HORAVERAO.CSV	Datas de início e fim do horário de verão
XX_AAAA-MM-DD_HORIZONTE.CSV	Informações sobre o horizonte de previsão
XX_AAAA-MM-DD_PATAMARES.CSV	Patamares de carga
XX_AAAA-MM-DD_SEPARADOR.CSV	Separador decimal
XX_AAAA-MM-DD_TEMPHIST.CSV	Dados históricos de temperatura equivalente em base horária
XX_AAAA-MM-DD_TEMPPREV.CSV	Previsões de temperatura equivalente em base horária

4.2.1 Histórico da carga global em base horária

Os dados de carga são disponibilizados em bases horárias e organizados em arquivos ".csv" com colunas separadas por ponto e vírgula. Cada média horária da carga deve ser acompanhado da respectiva data de ocorrência (ano/mês/dia/hora/minuto), conforme ilustrado na Figura 15.

	A	B	C	D	E	F
1	Ano	Mes	Dia	Hora	Min	Carga
2	2014	1	1	2	0	4880,6
3	2014	1	1	3	0	4829,4
4	2014	1	1	4	0	4824,5
5	2014	1	1	5	0	4811,9
6	2014	1	1	6	0	4758,4
7	2014	1	1	7	0	4681,2
8	2014	1	1	8	0	4474,4
9	2014	1	1	9	0	4425,1
10	2014	1	1	10	0	4397,3
11	2014	1	1	11	0	4365,6
12	2014	1	1	12	0	4368,4
13	2014	1	1	13	0	4386,1
14	2014	1	1	14	0	4379,3
15	2014	1	1	15	0	4425,5
16	2014	1	1	16	0	4500,4
17	2014	1	1	17	0	4538,6
18	2014	1	1	18	0	4507,3
19	2014	1	1	19	0	4478,7
20	2014	1	1	20	0	4734,7
21	2014	1	1	21	0	5162,1
22	2014	1	1	22	0	5284,9
23	2014	1	1	23	0	5314,9
24	2014	1	2	0	0	5321,9

Figura 15. Exemplo de arquivo XX_AAAA-MM-DD_CARGAHIST.CSV.

4.2.2 Pesos das previsões no cálculo da previsão combinada

Os pesos atribuídos às previsões obtidas pelos diferentes modelos disponibilizados no PrevcargaDESSEM são organizados em arquivos ".csv" com colunas separadas por vírgula, conforme ilustrado na Figura 16.

	A	B	C	D
1	Temperatura	Kernel	Modelo	Coeficientes
2	NA	NA	(Intercept)	167,0733007
3	3	1	Linear.TmpMaxMin	0,109184477
4	1	1	Linear.TmpMaxima	-0,75349304
5	2	1	Linear.TmpMedia	1,320650741
6	0	1	Linear.Univariado	0,372927521
7	1	0	Radial.TmpMaxima	0,664559545
8	2	0	Radial.TmpMedia	-1,329916001
9	0	0	Radial.Univariado	0,193426246
10	1	2	RedeNeural.TmpMaxima	0,229924509
11	2	2	RedeNeural.TmpMedia	0,093888001
12	0	2	RedeNeural.Univariado	0,086768775

Figura 16. Exemplo de arquivo XX_AAAA-MM-DD_COMBINA.CSV.

4.2.3 Lista de feriados e dias especiais

A tabela com informações de feriados e dias especiais deve conter a data de cada feriado (no formato ano/mês/dia) ao longo do período do histórico de carga e o tipo do feriado/dia especial, cuja codificação deve seguir a nomenclatura indicada na Figura 17a. Na Figura 17b tem-se um exemplo do arquivo XX_AAAA-MM-DD_FERIADOS.CSV com dados de feriados e dias especiais.

Tipo do dia	Código
Feriado Nacional perfil de sábado	1
Feriado Nacional perfil de domingo	2
Segunda-feira da semana do Carnaval	3
Terça-feira da semana do Carnaval	4
Quarta-feira de cinzas	5
Natal e 1º de Janeiro	6
Dia após Natal ou após 1º de Janeiro	7
Véspera de Natal ou véspera 1º de Janeiro	8
Véspera de feriado	9
Dia após feriado	10
Dia especial e recesso	11
Blecaute, Jogos da Copa, etc.	12

	A	B	C	D
1	Ano	Mes	Dia	Tipo
2	2014	1	1	6
3	2014	1	2	7
4	2014	2	28	9
5	2014	3	3	3
6	2014	3	4	4
7	2014	3	5	5
8	2014	4	17	9
9	2014	4	18	2
10	2014	4	20	2

a) Codificação de feriados e dias especiais

b) Arquivo XX_AAAA-MM-DD_FERIADOS.CSV

Figura 17. Arquivo XX_AAAA-MM-DD_FERIADOS.CSV com datas de feriados e dias especiais.

4.2.4 Horário de verão

As datas de início e fim do horário de verão devem ser informadas no arquivo XX_AAAA-MM-DD_HORAVERAO.CSV, cujo formato ano/mês/dia para cada edição do horário de verão ao longo do período do histórico de carga deve seguir o padrão indicado na Figura 18.

	A	B
1	Data.inicial	Data.final
2	20/10/2013	16/02/2014
3	19/10/2014	22/02/2015
4	18/10/2015	21/02/2016
5	16/10/2016	19/02/2017
6	15/10/2017	18/02/2018
7	04/11/2018	17/02/2019

Figura 18. Exemplo do arquivo XX_AAAA-MM-DD_HORAVERAO.CSV.

4.2.5 Dados do horizonte de previsão

As informações sobre o horizonte de previsão são disponibilizadas no arquivo XX_AAAA-MM-DD_HORIZONTE.CSV e incluem as datas dos instantes inicial e final do horizonte de previsão no formato ano/mês/dia/hora/minuto e as datas dos feriados no período de previsão no formato ano/mês/dia, acompanhadas da indicação do tipo do feriado. Na Figura19 tem-se um exemplo do arquivo XX_AAAA-MM-DD_HORIZONTE.CSV.

	A	B	C	D	E	F	G
1	Variável	Ano	Mês	Dia	Hora	Minuto	Feriado/Dia Especial
2	Instante inicial da previsão	2017	4	1	0	0	
3	Data final da previsão	2017	6	1	1	0	
4	Início do horário de verão	2016	10	16	0	0	
5	Fim do horário de verão	2017	2	19	0	0	
6	feriado	2017	4	14	0	0	0
7	feriado	2017	4	16	0	0	0
8	feriado	2017	4	21	0	0	0
9	feriado	2017	4	23	0	0	1

Figura 19. Exemplo do arquivo XX_AAAA-MM-DD_HORIZONTE.CSV.

4.2.6 Patamares de carga

O arquivo XX_AAAA-MM-DD_PATAMARES.CSV deve conter a associação das horas do dia aos patamares de carga, conforme ilustrado na Figura 20, cujos valores 1, 2, e 3 denotam os patamares de carga pesada, média e leve respectivamente, conforme indicado na Tabela 2.

	A	B	C	D	E	F	G
1	Hora	dia util inverno	fds inverno	dia util intermediario	fds intermediario	dia util verao	fds verao
2	00:00	3	3	3	3	3	3
3	01:00	3	3	3	3	3	3
4	02:00	3	3	3	3	3	3
5	03:00	3	3	3	3	3	3
6	04:00	3	3	3	3	3	3
7	05:00	3	3	3	3	3	3
8	06:00	3	3	3	3	3	3
9	07:00	2	3	3	3	3	3
10	08:00	2	3	2	3	2	3
11	09:00	2	3	2	3	2	3
12	10:00	1	3	1	3	1	3
13	11:00	1	3	1	3	1	3
14	12:00	1	3	1	3	1	3
15	13:00	1	3	1	3	1	3
16	14:00	1	3	1	3	1	3
17	15:00	1	3	1	3	1	3
18	16:00	1	3	1	3	1	3
19	17:00	1	3	1	3	1	3
20	18:00	1	2	1	2	2	3
21	19:00	1	2	1	2	2	3
22	20:00	1	2	2	2	2	2
23	21:00	1	2	2	2	2	2
24	22:00	2	3	2	3	2	2
25	23:00	2	3	2	3	2	3

Figura 20. Arquivo XX_AAAA-MM-DD_PATAMARES.CSV.

4.2.7 Separador decimal

O arquivo XX_AAAA-MM-DD_SEPARADOR.CSV deve conter o tipo de separador decimal usado nos dados, conforme indicado na Figura 21.

	A
1	x
2	,

Figura 21. Arquivo XX_AAAA-MM-DD_SEPARADOR.CSV.

4.2.8 Histórico das temperaturas horárias

O arquivo com dados históricos de temperatura deve conter os valores médios da temperatura equivalente (°C) em base horária. Adicionalmente, cada valor médio deve ser acompanhado da respectiva data de ocorrência, especificada em termos de ano/mês/dia/hora/minuto, conforme ilustrado na Figura 22.

	A	B	C	D	E	F
1	Ano	Mês	Dia	Hora	Min	Temperatura
2	2015	3	27	1	0	22
3	2015	3	27	2	0	23
4	2015	3	27	3	0	22
5	2015	3	27	4	0	23
6	2015	3	27	5	0	25

Figura 22. Exemplo do arquivo XX_AAAA-MM-DD_TEMP HIST.CSV.

4.2.9 Histórico das temperaturas horárias

O arquivo com previsões de temperatura deve conter os valores médios da temperatura equivalente (°C) em base horária ao longo do período de previsão. Cada linha do arquivo deve conter uma previsão da média horária de temperatura equivalente e sua respectiva data especificada como ano/mês/dia/hora/minuto, conforme ilustrado na Figura 23.

	A	B	C	D	E	F
1	Ano	Mês	Dia	Hora	Min	Temperatura
2	2017	4	1	1	0	22
3	2017	4	1	2	0	23
4	2017	4	1	3	0	22
5	2017	4	1	4	0	23
6	2017	4	1	5	0	25

Figura 23. Exemplo do arquivo XX_AAAA-MM-DD_TEMPPREV.CSV com previsões de temperatura equivalente.

4.2.10 Ponderação dos modelos de previsão

O arquivo XX_AAAA-MM-DD_COMBINA.CSV contém a ponderação dos modelos de previsão para cada intervalo semi-horário ao longo de um horizonte de previsão de 8 dias, i.e., 384 conjuntos de ponderação. Este arquivo deve ser informado no caso da execução de uma previsão combinada, cujo resultado é a soma ponderada das previsões obtidas por cada um dos 12 modelos de previsão disponibilizados pelo PrevcargaDESSEM, conforme a seguir:

- SVM com kernel linear & temperatura máxima
- SVM com kernel linear & temperaturas máxima e mínima
- SVM com kernel linear & temperatura média
- SVM com kernel linear & sem temperatura
- SVM com kernel radial & temperatura máxima
- SVM com kernel radial & temperaturas máxima e mínima
- SVM com kernel radial & temperatura média
- SVM com kernel radial & sem temperatura
- Rede neural artificial & temperatura máxima
- Rede neural artificial & temperaturas máxima e mínima
- Rede neural artificial & temperatura média
- Rede neural artificial & sem temperatura

Vale destacar que para cada intervalo semi-horário, o respectivo conjunto de ponderação define a constante de alisamento α e os pesos atribuídos aos previsores. A constante alfa determina pesos aos padrões entrada/saída e tem por finalidade capturar o efeito das idades dos padrões de treinamento no ajuste dos modelos, de tal forma que os padrões mais recentes recebam pesos maiores.

5. Execução em ambiente R

Conforme descrito na seção 4.2.10, o PrevcargaDESSEM disponibiliza 12 modelos de previsão. O programa tem dois modos de execução: individual e combinada. No modo de execução individual o usuário configura o programa para a execução de um dos 12 modelos disponíveis, enquanto no modo de previsão combinada o programa executa automaticamente os 12 modelos e ao final fornece a previsão resultante da combinação dos 12 modelos.

Para ilustrar a execução do PrevcargaDESSEM considere o *deck* de dados SECO_2020-07-29.zip, disponibilizado no Portal do Sintegre ONS e descompactado no diretório "c:/prevcarga".

5.1 Execução no modo individual

A seguir, tem-se o código para execução no modo individual, por exemplo, um modelo SVM com *kernel* radial e temperatura máxima como variável explicativa.

```
library(prevcargadessem)

# diretório onde estão localizados os arquivos de dados em formato csv
input="c:/prevcarga"

# diretório onde serão gravados os arquivos de saída
output="c:/prevcarga"

# formato do arquivo de dados: 0 para csv e 1 para data.frame
formato=0

# quatro opções de modelagem da temperatura diária
# 0 sem temperatura, 1 máxima, 2 média e 3 temperaturas mínima e máxima.
temp=1

# vetor formado por elementos binários, cada um referente a um dia do horizonte de previsão, 0 para
# não considerar a temperatura e 1 para considerar a temperatura.
xtemp=c(1,1,1,rep(0,5)) # considera a temperatura apenas nos três primeiros dias do horizonte

# três opções para o horizonte de previsão
# 0 horizonte variável em função do dia da previsão conforme Figura 1, 1 horizonte de oito dias,
# 2 horizonte até a data informada no arquivo XX_AAAA-MM-DD_HORIZONTE.CSV (seção 4.2.5)
horizonte=1

# número de dias iniciais com previsões semi-horárias
seminhora=3 # neste caso previsões por patamar a partir do quarto dia do horizonte

# método de previsão: 0 para SVM com kernel radial, 1 para SVM com kernel linear
# 2 para RNA com dois neurônios na camada intermediária.
kernel=0

# seleciona entre previsão individual (0) e previsão combinada (1)
combina=0

# acrônimo identificador do subsistema (N,NE,S,SECO)acompanhado da data
```



```
carga="SECO_2020-07-31"  
# constante de alisamento  
alfa=0
```

```
# executa a previsão de carga e guarda os resultados em prev, um objeto tipo lista  
prev=previsordessem(caminho=input,caminhosaida=output,tipoinput=formato,tipotemp=temp,tipotem  
pdia=xtemp,horizonteprevisao=horizonte,diasprevisaosemihora=semihora,tipokernel=kernel,tipocombi  
nada=combina, prefixo=carga,alfa=alfa)
```

Para maiores detalhes digite help(previsordessem) no console do R para abrir um texto explicativo dos argumentos da função previsordessem. A seguir, tem-se a saída apresentada no ambiente R durante a execução do programa.

```
CEPEL - Centro de Pesquisas de Energia Eletrica  
PPPP RRRR EEEEE V V CCCC AAA RRRR GGGG AAA DDDD EEEEE SSSSS SSSSS EEEEE M M  
P P R R E V V C A A R R G A A D D E S S E MM MM  
PPPP RRRR EEEEE V V C A A RRRR G GG A A D D EEEEE SSSSS SSSSS EEEEE M M M  
P R R E V V C AAAAA R R G G AAAAA D D E S S E M M  
P R R EEEEE V CCCC A A R R GGGG A A DDDD EEEEE SSSSS SSSSS EEEEE M M  
[1] "VERSAO 1.26.8"  
[1] "alfa= 0"  
[1] "Opcoes de execucao"  
[1] "Dados em arquivos CSV"  
[1] "Temperatura maxima no modelo de previsao da demanda media diaria"  
[1] "kernel radial no modelo SVM para previsao da demanda media diaria"  
[1] "previsao dos perfis dos dias especiais por semelhanca"  
[1] "Nome da serie: SECO_2020-07-31_"  
[1] "CONCLUIDA LEITURA DO ARQUIVO PARAMETROS.CSV"  
[1] "Mensagem: Horizonte de previsao informado 2020-08-07"  
[1] "Mensagem: Horizonte de previsao alterado para 2020-08-07"  
[1] "CONCLUIDA LEITURA DO ARQUIVO c:/prevcarga/SECO_2020-07-31_HORIZONTE.csv"  
[1] "CONCLUIDA LEITURA DO ARQUIVO c:/prevcarga/SECO_2020-07-31_CARGAHIST.csv"  
[1] "CONCLUIDA LEITURA DO ARQUIVO c:/prevcarga/SECO_2020-07-31_FERIADOS.csv"  
[1] "CONCLUIDA LEITURA DO ARQUIVO c:/prevcarga/SECO_2020-07-31_HORAVERAO.csv"  
[1] "CONCLUIDA LEITURA DO ARQUIVO c:/prevcarga/SECO_2020-07-31_TEMPHIST.csv"  
[1] "CONCLUIDA LEITURA DO ARQUIVO c:/prevcarga/SECO_2020-07-31_TEMPPREV.csv"  
[1] "CONCLUIDA LEITURA DO ARQUIVO COM HORARIOS DOS PATAMARES DE CARGA"  
[1] "CONCLUIDA LEITURA DO ARQUIVO DOS PESOS PARA A RODADA COMBINADA"  
[1] "MONTAGEM DA SERIE DE CARGA GLOBAL"  
[1] "TRATAMENTO DE DADOS"  
[1] "TREINA MODELOS E EXECUTA PREVISAO"  
[1] "Treina modelo 1 para previsao da demanda media diaria"  
[1] "Treina modelo 1 para previsao da demanda media diaria sem temperatura"  
[1] "Previsao para o dia 2020-07-31"  
[1] "Previsao para o dia 2020-08-01"  
[1] "Previsao para o dia 2020-08-02"  
[1] "Previsao para o dia 2020-08-03"  
[1] "Previsao para o dia 2020-08-04"  
[1] "Previsao para o dia 2020-08-05"  
[1] "Previsao para o dia 2020-08-06"  
[1] "Previsao para o dia 2020-08-07"  
[1] "CALCULA PREVISAO SEMI HORARIA"  
[1] "AGREGA PREVISAOES EM PATAMARES"  
[1] "GRAVANDO ARQUIVOS DE SAIDA"  
[1] "GRAVA ARQUIVO COM PREVISAOES NO FORMATO DESSEM"  
[1] "EXECUCAO CONCLUIDA COM SUCESSO"
```

Após a execução do PrevcargaDESSEM, as previsões de carga são armazenadas em objetos R *data.frames*, arquivos ".csv" e arquivo ".txt". Os arquivos ".csv" e ".txt" são gravados no diretório

informado em "caminhosaída" na chamada da função `previsordessem`. Mais especificamente, os seguintes resultados são exportados pelo programa:

- Previsões de carga global em MW médio em base semi-horária e patamar (arquivo `PREVCARGADESSEM_CARGA_GLOBAL.CSV` e `data.frame` `prev$CARGA_GLOBAL`).
- Previsões de carga global em MW médio em base semi-horária e patamar no formato DESSEM (arquivo `DP.txt` e `data.frame` `prev$SAIDADESSEM`).
- Previsões de carga global em MW médio em base horária (arquivo `PREVCARGADESSEM_CARGA_GLOBAL_HORARIA.CSV` e `data.frame` `prev$PREVHORARIA`).

As previsões da carga global em base semi-horária e por patamar são gravadas no arquivo `PREVCARGADESSEM_CARGA_GLOBAL.CSV`. Na Figura 24 tem-se um extrato do arquivo. O arquivo possui quatro colunas, a primeira mostra o dia e horário das previsões, a segunda coluna (MW Dessem) mostra a previsão de carga com resolução semi-horária nas primeiras 48 horas do horizonte (conforme definido na chave 4) e por patamar nas demais horas, a terceira coluna mostra as previsões em base semi-horária ao longo de todo o horizonte de previsão e a quarta coluna exibe as previsões por patamar, ao longo do horizonte. Para acessar os mesmos resultados em uma versão `data.frame` digite `prev$CARGA_GLOBAL`.

Um arquivo semelhante denominado `PREVCARGADESSEM_CARGA_GLOBAL_HORARIA.CSV`, mostra as previsões em base horária, cuja versão em formato `data.frame` pode ser acessada pelo comando `prev$PREVHORARIA`.

	A	B	C	D
1		MW Dessem	MW semi horaria	MW patamar
2	31/07/2020 00:30	31.213,30	31.213,30	28.866,57
3	31/07/2020 01:00	30.022,30	30.022,30	28.866,57
4	31/07/2020 01:30	29.098,92	29.098,92	28.866,57
5	31/07/2020 02:00	28.503,84	28.503,84	28.866,57
6	31/07/2020 02:30	28.186,60	28.186,60	28.866,57
7	31/07/2020 03:00	27.915,30	27.915,30	28.866,57
8	31/07/2020 03:30	27.653,86	27.653,86	28.866,57
9	31/07/2020 04:00	27.522,65	27.522,65	28.866,57
10	31/07/2020 04:30	27.578,66	27.578,66	28.866,57
11	31/07/2020 05:00	27.894,74	27.894,74	28.866,57
12	31/07/2020 05:30	28.462,34	28.462,34	28.866,57
13	31/07/2020 06:00	29.175,87	29.175,87	28.866,57
14	31/07/2020 06:30	29.992,48	29.992,48	28.866,57
15	31/07/2020 07:00	30.911,14	30.911,14	28.866,57
16	31/07/2020 07:30	31.896,45	31.896,45	36.268,91
17	31/07/2020 08:00	32.902,42	32.902,42	36.268,91
18	31/07/2020 08:30	33.853,36	33.853,36	36.268,91
19	31/07/2020 09:00	34.583,48	34.583,48	36.268,91
20	31/07/2020 09:30	35.097,72	35.097,72	36.268,91

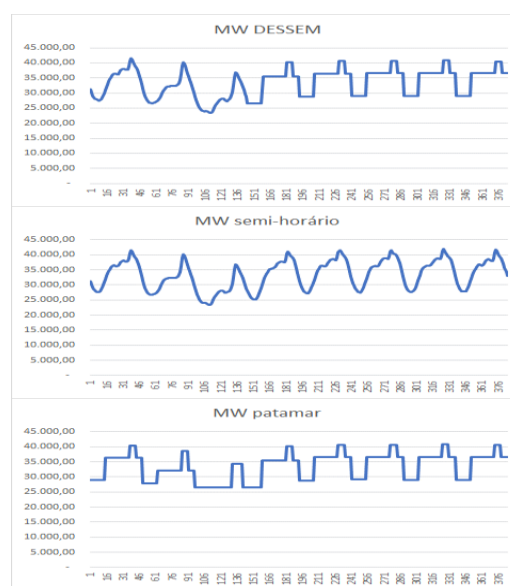


Figura 24. Arquivo `PREVCARGADESSEM_CARGA_GLOBAL.csv`.

Adicionalmente, o programa produz o arquivo `DP.TXT` ilustrado na Figura 25 com as previsões semi-horárias no formato para o DESSEM.

```

&
&
& CARGA
&
& ss di hi m df hf m Demanda
&X XX XX XX X XX XX X XXXXXXXXXXXX
DP 1 31 0 1 F 31213
DP 1 31 1 0 F 30022
DP 1 31 1 1 F 29099
DP 1 31 2 0 F 28504
DP 1 31 2 1 F 28187
DP 1 31 3 0 F 27915
DP 1 31 3 1 F 27654
DP 1 31 4 0 F 27523
DP 1 31 4 1 F 27579
DP 1 31 5 0 F 27895
DP 1 31 5 1 F 28462
DP 1 31 6 0 F 29176
DP 1 31 6 1 F 29992
DP 1 31 7 0 F 30911
DP 1 31 7 1 F 31896
DP 1 31 8 0 F 32902
DP 1 31 8 1 F 33853
DP 1 31 9 0 F 34583
DP 1 31 9 1 F 35098
DP 1 31 10 0 F 35618
DP 1 31 10 1 F 36137
DP 1 31 11 0 F 36433
DP 1 31 11 1 F 36470
DP 1 31 12 0 F 36290
DP 1 31 12 1 F 36059

```

O arquivo ao lado informa uma previsão de 31.213 MW para 00:30 h do dia 31, uma previsão de 30.022 MW para 01:00 h do mesmo dia, uma previsão de 29.099 MW para 01:30 h e assim sucessivamente.

Figura 25. Arquivo DP.TXT com previsões semi-horárias no formato DESSEM.

5.2 Execução no modo combinado

A seguir, tem-se o código para execução no modo combinado:

```
library(prevcargadessem)
```

```
# diretório onde estão localizados os arquivos de dados em formato csv
input="c:/prevcarga"
```

```
# diretório onde serão gravados os arquivos de saída
output="c:/prevcarga"
```

```
# formato do arquivo de dados: 0 para csv e 1 para data.frame
formato=0
```

```
# número de dias iniciais com previsões semi-horárias
seminhora=3 # neste caso previsões por patamar a partir do quarto dia do horizonte
```

```
# seleciona entre previsão individual (0) e previsão combinada (1)
combina=1
```

```
# acrônimo identificador do subsistema (N,NE,S,SECO)acompanhado da data
carga="SECO_2020-07-31"
```

```
# constante de alisamento
alfa=0
```

```
# executa a previsão de carga e guarda os resultados em prev, um objeto tipo lista
prevcomb=previsordessem(caminho=input,caminhosaida=output,tipoinput=formato,diasprevisaoseminhora=seminhora, tipocombinada=combina,prefixo=carga,alfa=alfa)
```

Após a execução do `PrevcargaDESSEM`, as previsões de carga são armazenadas em objetos R `data.frames`, arquivos ".csv" e arquivo ".txt". Os arquivos ".csv" e ".txt" são gravados no diretório informado em "caminhosaida" na chamada da função `previsordessem`. Mais especificamente, os seguintes resultados são exportados pelo programa:

- Previsões de carga global em MW médio em base semi-horária e patamar (arquivo `PREVCARGADESSEM_CARGA_GLOBAL.CSV` e `data.frame prevcomb$CARGA_GLOBAL`).
- Previsões de carga global em MW médio em base semi-horária e patamar no formato DESSEM (arquivo `DP.txt` e `data.frame prevcomb$SAIDADESSEM`).

As previsões da carga global em base semi-horária e por patamar são gravadas no arquivo `PREVCARGADESSEM_CARGA_GLOBAL.CSV` no mesmo formato ilustrado na Figura 24. Para acessar os mesmos resultados em uma versão `data.frame` digite `prevcomb$CARGA_GLOBAL`.

Referências bibliográficas

Baghel, M.; Ghosh, A.; Singh, N.K.; Singh, A.K. Short -Term Electric Load Forecasting Using SVR Implementing LibSVM Package and Python Code, Electrical, Computer and Electronics Engineering (UPCON), 2016 IEEE Uttar Pradesh Section International Conference on, 9-11 December, 2016.

Bates, J. M.; Granger, C. W. J. The combination of forecasts. *Operations Research Quarterly*, 20:451–468, 1969.

Bozic, M; Stojonavic, M. Application of SVM Methods for Mid-Term Load Forecasting, *Serbian Journal of Electrical Engineering*, Vol. 8, No. 1, pp. 73-83, February 2011.

Cannon, R.L.; Dav e, J.V.; Bezdek, J.C. Efficient Implementation of the Fuzzy c-Means Clustering Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, n. 2, pp. 248-255, March, 1986.

Chen, B.J.; Chang, M.W.; Lin, C.J. Load forecasting using support vector machines: a study on EUNITE Competition 2001, *IEEE Transactions on power systems*, v. 19, n. 4, November 2004.

Ceperic, E; Ceperic, V; Baric, A. A Strategy for Short-Term Load Forecasting by Support Vector Regression Machines, *IEEE Transactions on Power Systems*, Volume: 28, Issue: 4, pages 4356 – 4364, November 2013.

Domingos, P. O Algoritmo Mestre, São Paulo: Novatec, 2017

Dubinsky, A.; Elperin, T. A method for calculating a load curve using average values of load over time intervals, *International Journal of Electrical Power & Energy Systems*, v. 19, n. 6, pp. 393-395, August, 1997.

Grus, J. Data science do zero: primeiras regras com o Python, Rio de Janeiro: Alta Books, 2016.

Hamel, L. Knowledge discovery with support vector machines, Wiley, 2009.

Haykin, S. Neural networks and learning machines, Third Upper Saddle River, NJ: Pearson Education, 2009.

Hippert, H.S.; Pedreira, C.E.; Souza, R.C. Neural Networks for Short-Term Load Forecasting: A review and evaluation, IEEE Transactions on Power Systems, v. 16, n. 1, February, 2001.

Hong, W.C. Electric load forecasting by support vector model, Applied Mathematical Modelling Volume 33, Issue 5, Pages 2444-2454, May 2009.

Kodogiannis, V.S.; Anagnostakis, E.M. A study of advanced learning algorithms for short-term load forecasting, Engineering Applications of Artificial Intelligence 12, pp. 159 - 173, 1999.

Kovács, Z.L. Redes Neurais Artificiais: fundamentos e aplicações, São Paulo: Edição Acadêmica, 1996.

Liao, G.C.; Tsao, T.P. Application of fuzzy neural networks and artificial intelligence for load forecasting, Electric Power Systems Research 70, pp. 237-244, 2004.

Martinez, W.L.; Martinez, A.R. Computational statistics handbook with matlab, Chapman & Hall/CRC, 2002.

Mastorocostas, P.A.; Theocharis, J.B.; Bakirtzis, A.G. Fuzzy modeling for short-term load forecasting using the orthogonal least squares method, IEEE, Transactions on Power Systems, vol. 14, n.1, february, 1999.

Niu, D.X.; Wang, Y.L. Support Vector Machines Based on Data Mining Technology in Power Load Forecasting, International Conference on Wireless Communications, Networking and Mobile Computing, Shanghai, China, September, 2007.

Park, H.Y.; Lee, B.H.; Son, J.H.; Ahn, H.S.; A comparison of neural network-based methods for load forecasting with selected input candidates, PowerTech, 2017 IEEE Manchester, 18-22 Jun , 2017.

Pessanha, J.F.M.; Justino, T.C. Uma Metodologia para Previsão de Carga no Curto-Prazo. XIII Simpósio de Especialistas em Planejamento da Operação e Expansão Elétrica, Foz do Iguaçu, 2014.

Pessanha, J.F.; Xavier, V.L.; Amaral, M.R.S.; Laurencel, L.C. Construindo tipologias de curvas de carga com o programa R, Pesquisa Operacional para o Desenvolvimento, v. 7, n. 1, 2015.

Pessanha, J.F.M.; Melo, A.C.G.; Justino, T.C.; Maceira, M.E.P. Combining Statistical Clustering Techniques and Exploratory Data Analysis to Compute Typical Daily Load Profiles - Application to the Expansion and Operational Planning in Brazil, IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 24-28 June, Boise-ID, USA, 2018.

R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2017

Smola, A.J; Scholkopf, B. A Tutorial on Support Vector Regression, Journal Statistics and Computing, Volume 14 Issue 3, Pages 199-222, August 2004.

Turkay, B.E.; Demren, D. Electrical load forecasting using support vector machines, Electrical and Electronics Engineering (ELECO), 2011 7th International Conference on, December 2011.

Vapnik, V.N.; Boser, E.B.; Guyon, I.M.; Training Algorithm for Optimal Margin Classifiers, Proceedings of the fifth annual workshop on Computational learning theory, Pages 144-152, Pennsylvania, July 1992.

Wang, J.H.; Wu, L.; Lu, H.Y. Special periods peak load analysis and superior forecasting method based on LS-SVM, International Conference on Wavelet Analysis and Pattern Recognition, Beijing, China, 2-4 November. 2007.

Werbos, P.J. Backpropagation through time: What it does and how to do it, Proceedings of the IEEE, v. 78, n. 10, October, 1990.

Xu, H.; Wang, J.H.; Zheng, S.Q. Online daily load forecasting based on support vector machines, Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005

Yadav, S.; Shukla, S. Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification, Advanced Computing (IACC), 2016 IEEE 6th International Conference on, February 2016.

Zhang, G.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art, International Journal of Forecasting, 14, pp. 35-62, 1998.