

Relatório de Previsão de Séries Temporais

Victor Carreira

Resumo—Este relatório apresenta o resultado da previsão realizada por uma rede neural artificial *Multi Layer Perceptron* do tipo *Back Propagation* em um banco de dados da concessionária de energia elétrica, Light, com o intuito de se fazer a previsão mensal da sensação térmica das diversas regiões ou unidades geográficas que fazem parte da sua área de abrangência. A análise de 9 experimentos realizados indicam que o melhor resultado da rede apresentou erros de MAPE de 3.3669 para o conjunto de validação e 4.3468 para o conjunto de treinamento. E valores de RMSE de 1.5717 para o conjunto de validação e 2.3305 para o conjunto de treinamento.

I. INTRODUÇÃO

As Redes Neurais Artificiais (RNA) são inspiradas em modelos sensoriais do processamento de tarefas realizadas pelo cérebro (Hagan et al., 1996). Uma RNA, portanto pode ser criada através da aplicação de algoritmos matemáticos que imitem a tarefa realizada por um neurônio (Nedjah et al., 2016). Uma rede neural artificial possui semelhanças com a rede biológica presente no sistema nervoso central, neste o cômputo de informações realizado do cérebro é feito através de uma vasta quantidade de neurônios interconectados (Feldman et al., 1988; Poulton, 2002). A comunicação entre essas células é realizada através de impulsos elétricos. Estes são transmitidos e recebidos por meio de sinapses nervosas entre axônios e dendritos. As sinapses são estruturas elementares e uma unidade funcional localizada entre dois neurônios (Krogh, 2008).

McCulloch and Pitts (1943) redigem o trabalho pioneiro onde foi modelado um neurônio cuja resposta dependia do *input*¹ que provinha de outros neurônios e do peso utilizado. Já Rosenblatt (1962) cria a teoria de convergência do *Perceptron* onde ele prova que modelos de neurônios possuem propriedades similares ao cérebro humano (Kanal, 2001). Neste sentido as rede neuronais artificiais podem realizar performances sofisticadas no reconhecimento de padrões, mesmo se alguns neurônios forem destruídos (Levy, 1997). Minsky and Papert (1969) demonstraram que um único *Perceptron* somente resolve uma classe muito limitada de problemas que podem ser linearizados.

A retropropagação é uma técnica específica para implementar a decida do gradiente no espaço de pesos para um rede de múltiplas camadas alimentada. A ideia básica é calcular as derivadas parciais de uma função aproximativa $F(\mathbf{w}, \mathbf{x})$ realizada pela rede em relação a todos os elementos do vetor de ajuste de pesos \mathbf{w} para um dado valor de entrada \mathbf{x} (Haykin, 1999).

Neste relatório são apresentados os resultados do Trabalho 02 MLP de Previsão de Séries Temporais da disciplina ELE 2394, Redes Neurais I, da Engenharia elétrica.

Professora: Marley. Eng. Elétrica. PUC-RIO

¹Valor de entrada

II. OBJETIVO E METODOLOGIA

As concessionárias de energia elétrica compram energia baseada na demanda futura. É notório que o clima (temperatura, precipitação, umidade relativa do ar, vento, etc) afeta o consumo de energia elétrica, principalmente o consumo de energia elétrica das classes residencial e comercial. A Light é uma concessionária caracterizada pelo consumo dessas duas classes. Esta característica fica evidente no caso de algumas regiões da Light, onde o percentual de clientes residenciais e comerciais é maior do que 90%.

A Light está situada no estado do Rio de Janeiro, e por isso apresenta grandes diversidades climáticas dentro da sua área de concessão devido à proximidade do mar e à presença de montanhas. Portanto, estudar os efeitos do clima em relação à energia faturada e à carga é fundamental para aumentar a eficácia na previsão do volume de consumo e, consequentemente, o volume de compra de energia. Assim, uma das metas a serem atingidas pelas concessionárias é reduzir os erros nas estimativas da carga, reduzindo os prejuízos na compra de energia. Dessa forma, a empresa poderá reduzir o risco no fluxo de caixa permitindo futuros investimentos na expansão e qualidade dos serviços prestados a seus clientes.

Portanto, o objetivo deste trabalho é desenvolver um modelo para a previsão mensal da sensação térmica das diversas regiões ou unidades geográficas que fazem parte da área de concessão da Light.

A partir do modelo de previsão da sensação térmica se poderá desenvolver um modelo original de previsão de carga e faturamento da empresa, onde além de agregar a previsão da sensação térmica, também considerará dados históricos de carga e faturamento.

Detalhamento

- Aplicação: Previsão mensal da Sensação Térmica em 8 regiões distintas do Município do Rio de Janeiro, parte da área de concessão da Light S.A.;
- Dados: Janeiro de 1998 a Dezembro de 2008;
- Modelo Neural: Treinamento de MLP com treinamento por Back Propagation;
- Saída da Rede: Previsão da demanda mensal 12 meses a frente (previsão multi-step);
- Treinamento e Validação: Valores entre Janeiro/1998 e Dezembro/2008;
- Teste: Janeiro a Dezembro de 2009;
- Software: MATLAB 2008b em diante.

Efetuar a previsão, 12 passos a frente, da sensação térmica mensal em uma das 8 regiões do Município do Rio de Janeiro (Centro, Zona Sul 1, Zona Sul 2, Zona Norte 1, Zona Norte 2, Oeste 1, Oeste 2, Oeste 3), variando os seguintes parâmetros de configuração e treinamento da Rede Neural.

Testes

- Janela: Variar o tamanho da janela de entrada, tentando obter o melhor desempenho de generalização;
- Codificação do Mês: Variar o método de codificação da informação sobre o mês da previsão – uma entrada real, 12 entradas binárias (uma para cada mês) e 4 entradas binárias codificadas;
- Topologia: Variar o número de processadores na camada escondida – escolher dois valores diferentes para cada configuração de janela de entrada; variar o tipo de função de ativação nas camadas escondida (logsig e tansig) e de saída (logsig e purelin);
- Treinamento: Utilizar o método de parada antecipada (early stop).

No final são apresentados os gráficos de treinamento e teste para cada configuração e o desempenho em termos das seguintes métricas:

- 1) MAPE (“Mean Absolute Percentage Error”);
- 2) RMSE (“Root Mean Square Error”).

O relatório apresenta uma tabela comparativa dessas duas métricas para todas as configurações de redes utilizadas. O relatório deve também conter uma discussão dos resultados encontrados em função dos parâmetros utilizados, tais como as diferentes codificações, os tamanhos da janela, os números de processadores e as condições de parada do treinamento com normalização dos dados de entrada.

III. PRINCÍPIO TEÓRICO

O neurônio de McCulloch and Pitts (1943) propõe um limite binário para a criação de um modelo. Este neurônio artificial registra uma soma de pesos de n sinais de entrada, x_j , $j = 1, 2, 3, \dots, n$, e fornece um *output*² de 1 caso esta soma esteja acima do limite u . Caso contrário o *output* é 0. Matematicamente essa relação pode ser descrita de acordo com a Eq. 1:

$$y = \theta \left(\sum_{j=1}^n w_j x_j - u \right) \quad (1)$$

Onde θ é o passo dado na posição 0, w_j é chamada sinapse-peso associado a um *jesimo* *input*. A título de simplificação a função limite³ u é considerada um outro peso $w_0 = -u$ anexado a um neurônio com um *input* constante $x_0 = 1$. Pesos positivos correspondem a uma sinapse **excitatória**, enquanto pesos negativos correspondem a uma sinapse **inibitória**. Este modelo contém uma série de simplificações que não refletem o verdadeiro comportamento dos neurônios biológicos (Mao, 1996).

Derivações do neurônio de McCulloch and Pitts (1943) na escolha das funções de ativação. Uma função largamente utilizada é a função sigmóide, que exibe uma suavização dos *outputs* a medida que o valor da função diminui (Mao, 1996;

Misra and Saha, 2010). Essa função de ativação pode ser expressa de acordo com a Eq. 2:

$$g(x) = 1/(1 + e^{-\beta x}) \quad (2)$$

Onde β é o parâmetro de inclinação.

As redes alimentadas diretamente são aquelas redes cujos grafos orientados se distinguem pela presença de um ou mais camadas ocultas e cujos nós são chamados de neurônios ocultos. A função do neurônio é intervir entre a camada externa e a saída da rede de maneira útil. Adicionando-se camadas ocultas a rede torna-se capaz de realizar estatísticas de ordem elevada (Haykin, 1999).

IV. RESULTADOS

A região escolhida para a realização deste trabalho foi a 1 localizada no centro do Rio de Janeiro. E o número de redes a serem testadas permaneceu 1 para todos os testes. Foram realizados um total de 10 diferentes testes.

O janelamento utilizado no primeiro experimento foi de 12, com um codificação do mês real, uma topologia composta de 13 processadores, o algoritmo de treinamento “trainlm”, com a função “tansig”, e com uma função de saída linear, e sem a utilização do critério de parada antecipada. A Fig. 1 apresenta a topologia da rede no primeiro experimento.

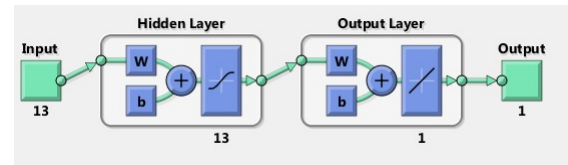


Figura 1: Topologia da rede para o primeiro teste realizado.

O conjunto de treinamento, validação e teste para o primeiro experimento é apresentado na Fig. 2.

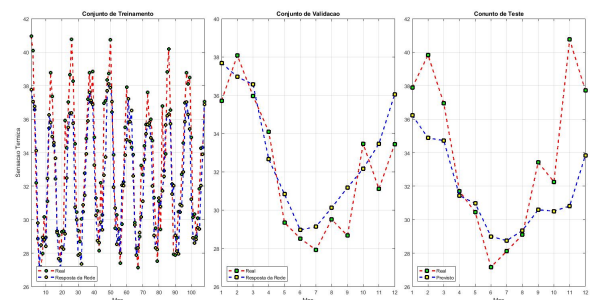


Figura 2: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros. MAPE de 4.2542 para o conjunto de validação e 5.8434 para

²Valor de saída

³Genericamente chamada de função de ativação

o conjunto de treinamento. E valores de RMSE de 1.6624 para o conjunto de validação e 2.7465 para o conjunto de treinamento.

O janelamento utilizado no segundo experimento foi de 6, com um codificação do mês real, uma topologia composta de 13 processadores, o algoritmo de treinamento "trainlm", com a função "tansig", e com uma função de saída linear, e sem a utilização do critério de parada antecipada. A Fig. 3 apresenta à topologia da rede no segundo experimento.

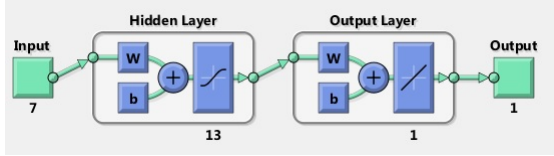


Figura 3: Topologia da rede para o segundo teste realizado.

O conjunto de treinamento, validação e teste para o segundo experimento é apresentado na Fig. 4.

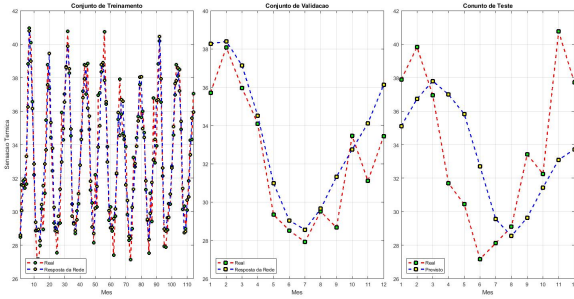


Figura 4: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros. MAPE de 4.7637 para o conjunto de validação e 3.9691 para o conjunto de treinamento. E valores de RMSE de 2.0730 para o conjunto de validação e 2.1397 para o conjunto de treinamento.

O janelamento utilizado no terceiro experimento foi de 6, com um codificação do mês 4 bits, uma topologia composta de 13 processadores, o algoritmo de treinamento "trainlm", com a função "tansig", e com uma função de saída linear, e sem a utilização do critério de parada antecipada. A Fig. 5 apresenta à topologia da rede no terceiro experimento.

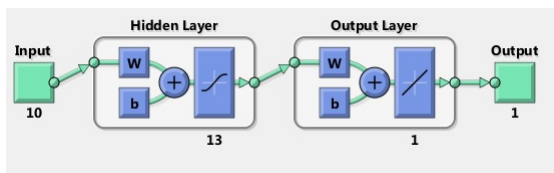


Figura 5: Topologia da rede para o terceiro teste realizado.

O conjunto de treinamento, validação e teste para o terceiro experimento é apresentado na Fig. 6.

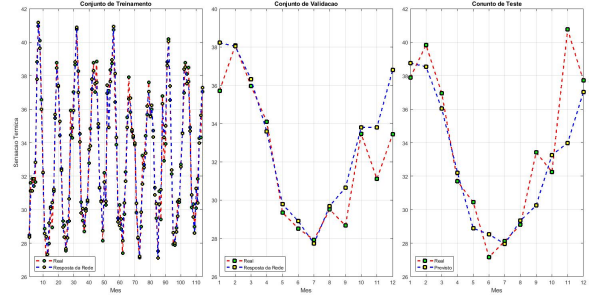


Figura 6: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros. MAPE de 3.3669 para o conjunto de validação e 4.3468 para o conjunto de treinamento. E valores de RMSE de 1.5717 para o conjunto de validação e 2.3305 para o conjunto de treinamento.

O janelamento utilizado no quarto experimento foi de 6, com um codificação do mês 12 bits, uma topologia composta de 13 processadores, o algoritmo de treinamento "trainlm", com a função "tansig", e com uma função de saída linear, e sem a utilização do critério de parada antecipada. A Fig. 7 apresenta à topologia da rede no quarto experimento.

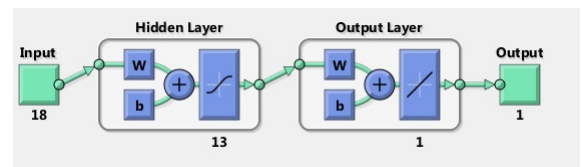


Figura 7: Topologia da rede para o quarto teste realizado.

O conjunto de treinamento, validação e teste para o quarto experimento é apresentado na Fig. 6.

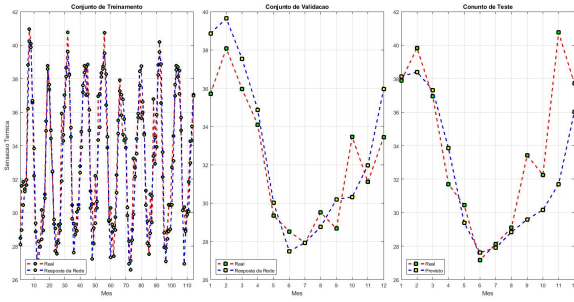


Figura 8: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros. MAPE de 4.4237 para o conjunto de validação e 5.2831 para o conjunto de treinamento. E valores de RMSE de 1.7508 para o conjunto de validação e 3.0548 para o conjunto de treinamento.

O janelamento utilizado no quinto experimento foi de 6, com um codificação do mês 12 bits, uma topologia composta de 3 processadores, o algoritmo de treinamento "trainlm", com a função "tansig", e com uma função de saída linear, e sem a utilização do critério de parada antecipada. A Fig. 9 apresenta à topologia da rede no quinto experimento.

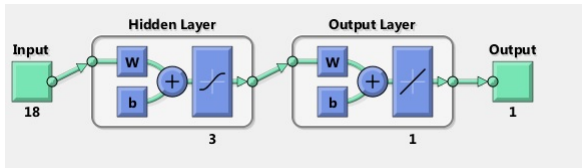


Figura 9: Topologia da rede para o quinto teste realizado.

O conjunto de treinamento, validação e teste para o quinto experimento é apresentado na Fig. 10.

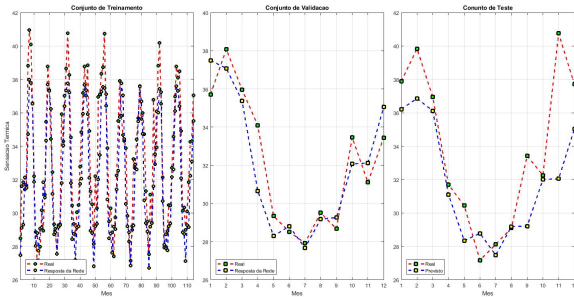


Figura 10: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros.

MAPE de 3.3555 para o conjunto de validação e 6.1024 para o conjunto de treinamento. E valores de RMSE de 1.4032 para o conjunto de validação e 3.1737 para o conjunto de treinamento.

O janelamento utilizado no sexto experimento foi de 6, com um codificação do mês 12 bits, uma topologia composta de 23 processadores, o algoritmo de treinamento "trainlm", com a função "tansig", e com uma função de saída linear, e sem a utilização do critério de parada antecipada. A Fig. 11 apresenta à topologia da rede no sexto experimento.

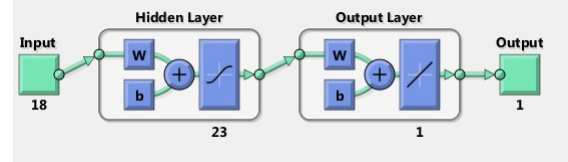


Figura 11: Topologia da rede para o sexto teste realizado.

O conjunto de treinamento, validação e teste para o sexto experimento é apresentado na Fig. 12.

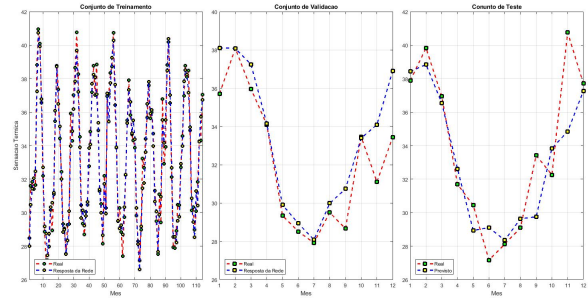


Figura 12: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros. MAPE de 3.6180 para o conjunto de validação e 4.5186 para o conjunto de treinamento. E valores de RMSE de 1.6659 para o conjunto de validação e 2.2323 para o conjunto de treinamento.

O janelamento utilizado no sétimo experimento foi de 6, com um codificação do mês 12 bits, uma topologia composta de 23 processadores, o algoritmo de treinamento "trainlm", com a função "logsig", e com uma função de saída linear, e sem a utilização do critério de parada antecipada. A Fig. 13 apresenta à topologia da rede no sétimo experimento.

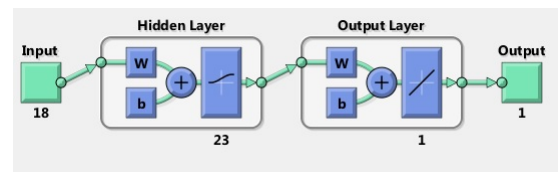


Figura 13: Topologia da rede para o sétimo teste realizado.

O conjunto de treinamento, validação e teste para o sétimo experimento é apresentado na Fig. 14.

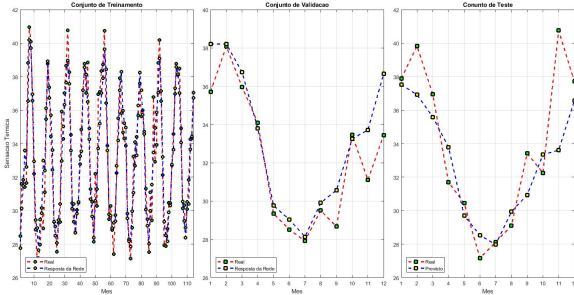


Figura 14: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros. MAPE de 3.4132 para o conjunto de validação e 5.0634 para o conjunto de treinamento. E valores de RMSE de 1.5364 para o conjunto de validação e 2.5437 para o conjunto de treinamento.

O janelamento utilizado no oitavo experimento foi de 6, com um codificação do mês 12 bits, uma topologia composta de 23 processadores, o algoritmo de treinamento "trainlm", com a função "logsig", e com uma função de saída "logsig", e sem a utilização do critério de parada antecipada. A Fig. 15 apresenta a topologia da rede no oitavo experimento.

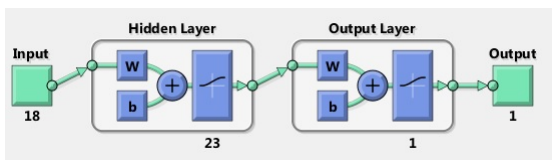


Figura 15: Topologia da rede para o oitavo teste realizado.

O conjunto de treinamento, validação e teste para o oitavo experimento é apresentado na Fig. 16.

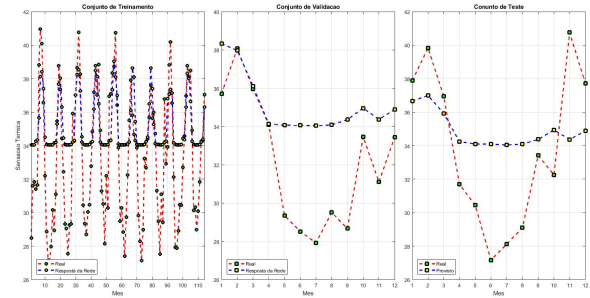


Figura 16: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros. MAPE de 10.0192 para o conjunto de validação e 10.9309 para o conjunto de treinamento. E valores de RMSE de 3.7165 para o conjunto de validação e 4.0335 para o conjunto de treinamento.

O janelamento utilizado no nono experimento foi de 6, com um codificação do mês 12 bits, uma topologia composta de 23 processadores, o algoritmo de treinamento "trainlm", com a função "logsig", e com uma função de saída "logsig", e com a utilização do critério de parada antecipada. A Fig. 17 apresenta a topologia da rede no nono experimento.

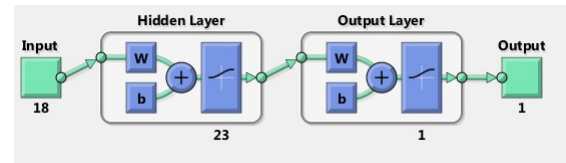


Figura 17: Topologia da rede para o nono teste realizado.

O conjunto de treinamento, validação e teste para o nono experimento é apresentado na Fig. 18.

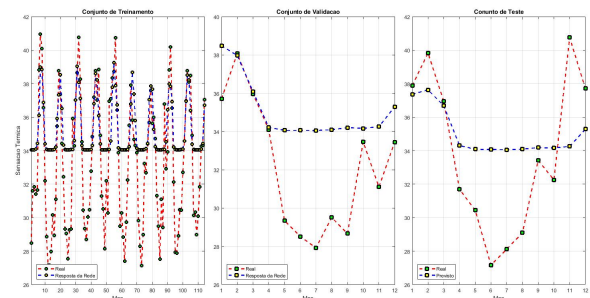


Figura 18: Os gráficos apresentam da esquerda para a direita os conjuntos de treinamento, validação e teste da rede para o primeiro experimento. Em vermelho é registrado o dado real, enquanto em azul é apresentado o ajuste da rede. Aonde o eixo cartesiano horizontal representa os meses e o eixo vertical a sensação térmica.

O presente teste apresentou os seguintes valores de erros.

MAPE de 9.8988 para o conjunto de validação e 10.1712 para o conjunto de treinamento. E valores de RMSE de 3.6942 para o conjunto de validação e 3.9312 para o conjunto de treinamento.

No final deste relatório é apresentado um quadro geral de todos os testes realizados.

V. CONCLUSÕES

Os experimentos realizados apresentam uma sequência lógica de alterações de parâmetros da rede. Isto é, somente um parâmetro é alterado por vez, enquanto os demais permanecem inalterados.

Dentre esta sequência de experimentos o que apresentou o melhor resultado foi experimento de número três, cujo janelamento utilizado foi o de 6, com um codificação do mês 4 bits, uma topologia composta de 13 processadores, o algoritmo de treinamento "trainlm", com a função "tansig", e com uma função de saída linear, e sem a utilização do critério de parada antecipada. Este apresentou erros de MAPE de 3.3669 para o conjunto de validação e 4.3468 para o conjunto de treinamento. E valores de RMSE de 1.5717 para o conjunto de validação e 2.3305 para o conjunto de treinamento.

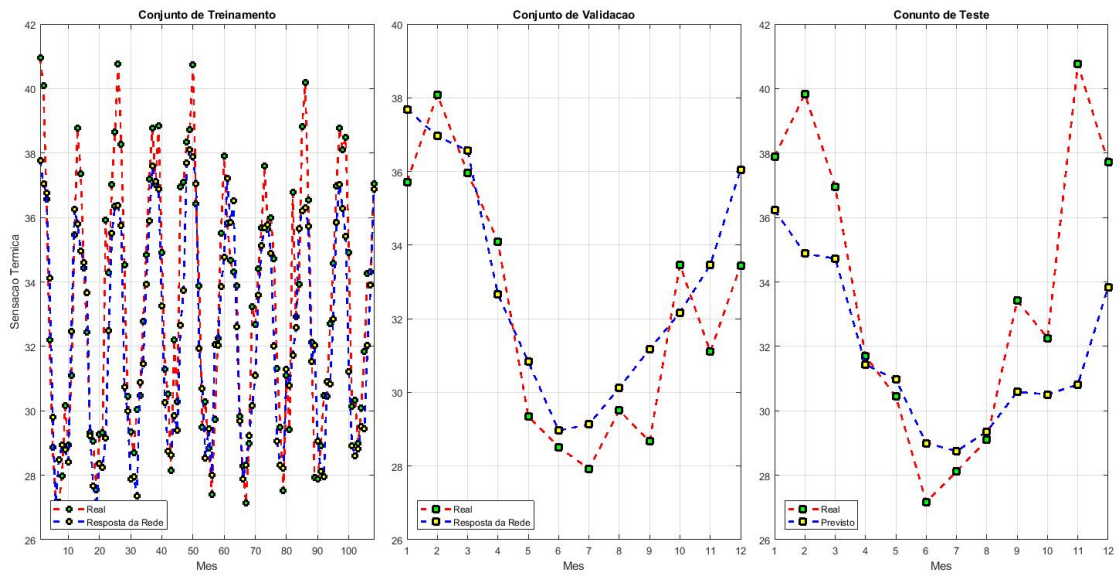
Em contrapartida, o experimento que apresentou o pior desempenho foi o oitavo cujo janelamento utilizado foi de 6, com um codificação do mês 12 bits, uma topologia composta de 23 processadores, o algoritmo de treinamento "trainlm", com a função "logsig", e com uma função de saída "logsig", e sem a utilização do critério de parada antecipada. Apresentando erros de MAPE de 10.0192 para o conjunto de validação e 10.9309 para o conjunto de treinamento. E valores de RMSE de 3.7165 para o conjunto de validação e 4.0335 para o conjunto de treinamento.

Essa discrepância registrada se deveu principalmente ao fato de que o experimento 3 possui uma janela de previsão menor do que o experimento 8. Outro fator importante para o aumento dos erros foi a utilização do critério de parada antecipada no experimento 8. Como neste caso o número de processadores foi aumentado para 23 seria necessário aumentar o número de ciclos de treinamento ou diminuir o critério de parada.

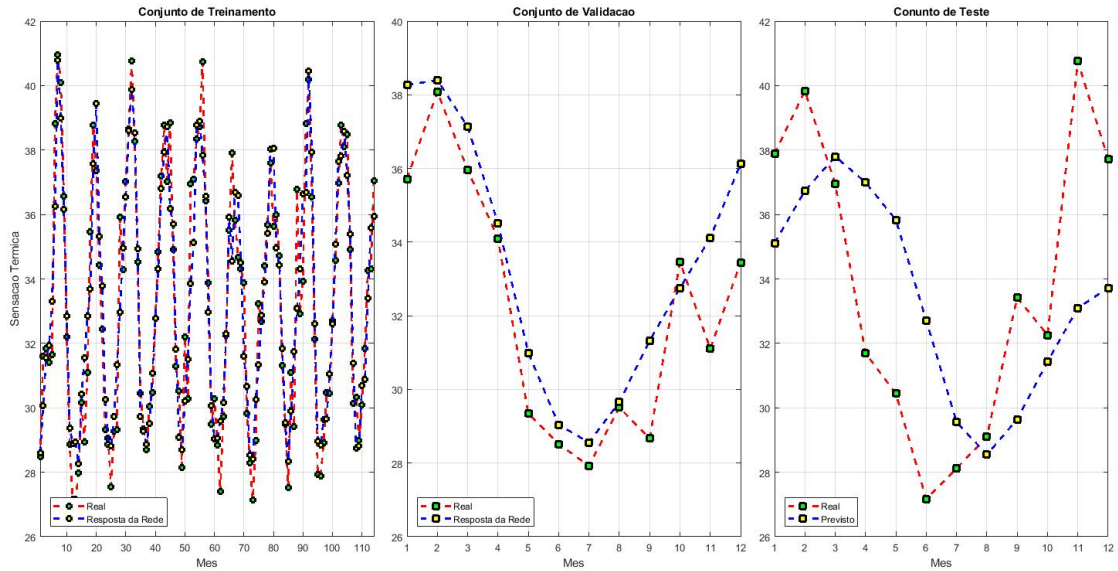
Outra questão importante foi a mudança da função de saída de linear para "logsig". A função linear possui um comportamento sem grandes variações entre os pontos que a descrevem. Enquanto que a função "logsig" já apresenta um comportamento mais variável. Soma-se a isso ao fato de a natureza do dado escolhido (centro do Rio de Janeiro) apresentar pouca variação no comportamento de temperatura ao longo dos meses se comparado com outras regiões da cidade.

REFERÊNCIAS

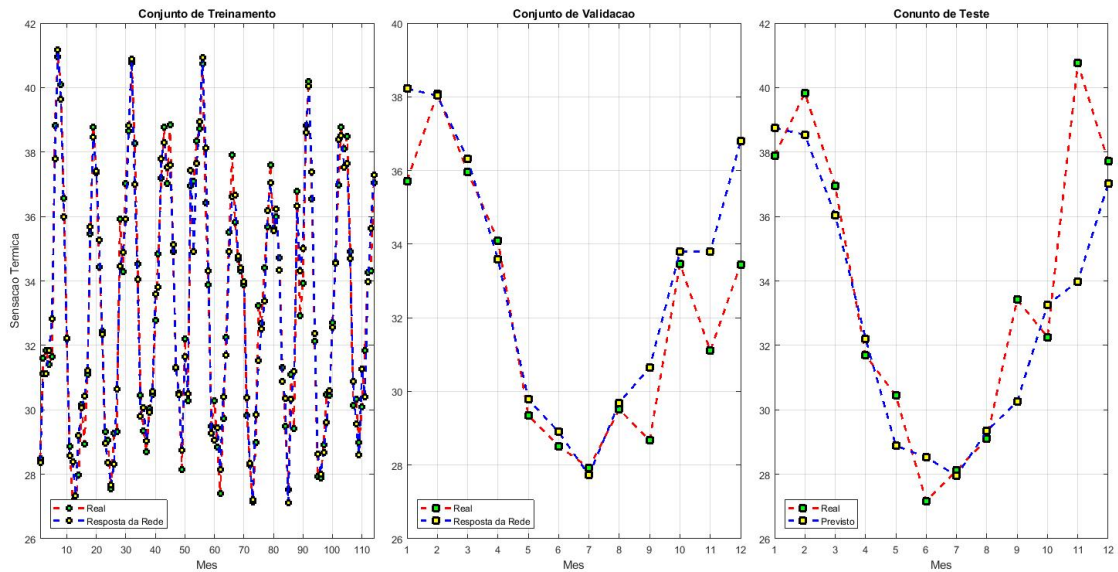
- Feldman, J. A., Fandy, M. A., and Goddard, N. H. (1988). Computing With Structured Neural Networks. *Computer*, 21(3):91–103.
- Hagan, M. T., Demuth, H. B., Beale, M. H., and De Jesús, O. (1996). *Neural Network Design*. page 1012.
- Haykin, S. (1999). *Neural Networks: a comprehensive foundation*. Prentice Hall Inc., Porto Alegre, 2 edition.
- Kanal, L. N. (2001). Perceptrons. *Encyclopedia of Computer Science*, pages 11215–11218.
- Krogh, A. (2008). What are artificial neural networks? *Nature biotechnology*, 26(2):195–197.
- Levy, S. (1997). The Computer. *Newsweek*, 130(22):28.
- Mao, J. (1996). Why artificial neural networks? *Communications*, 29:31–44.
- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133.
- Minsky, M. and Papert, S. (1969). *Perceptrons*. 2 edition.
- Misra, J. and Saha, I. (2010). Artificial neural networks in hardware: A survey of two decades of progress. *Neurocomputing*, 74(1-3):239–255.
- Nedjah, N., da Silva, F. P., de Sá, A. O., Mourelle, L. M., and Bonilla, D. A. (2016). A massively parallel pipelined reconfigurable design for M-PLN based neural networks for efficient image classification. *Neurocomputing*, 183:39–55.
- Poulton, M. M. (2002). Neural networks as an intelligence amplification tool: A review of applications. *Geophysics*, 67(3):979.
- Rosenblatt, F. (1962). Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms. *Archives of General Psychiatry*, 7:218–219.



(a) Experimento 1

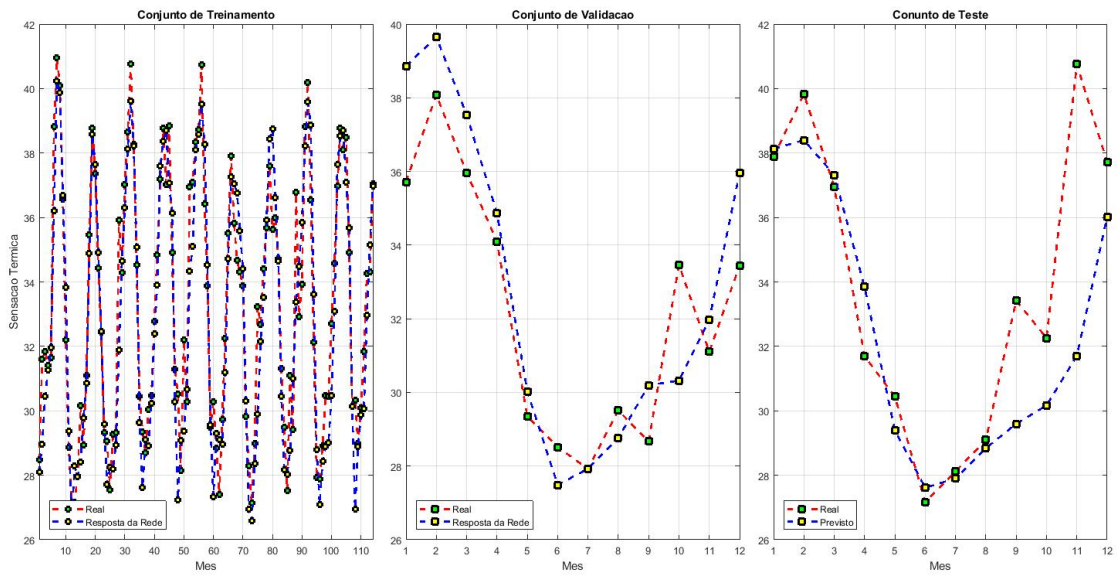


(b) Experimento 2

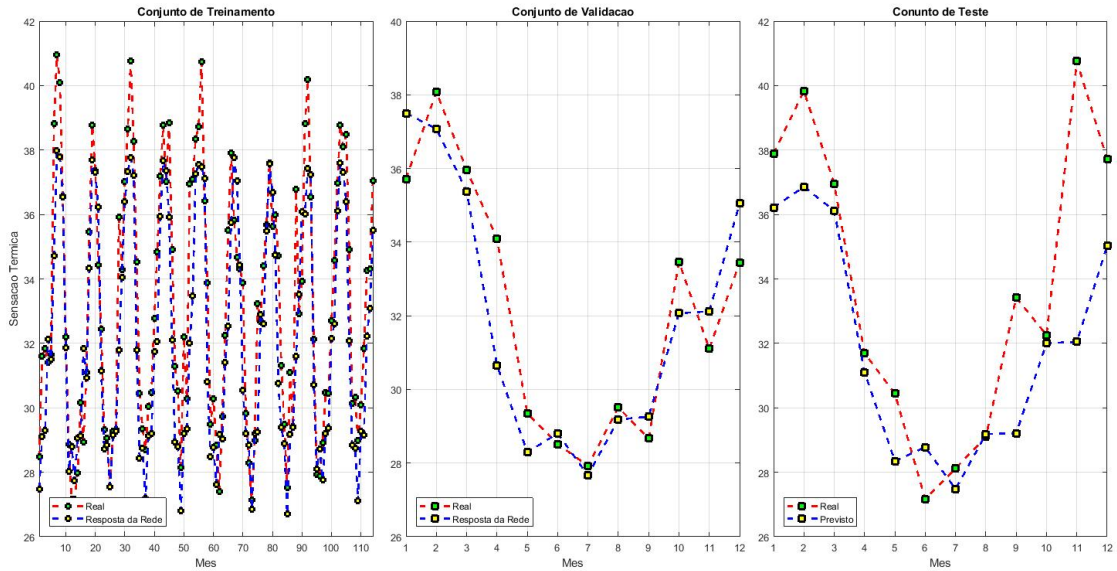


(c) Experimento 3

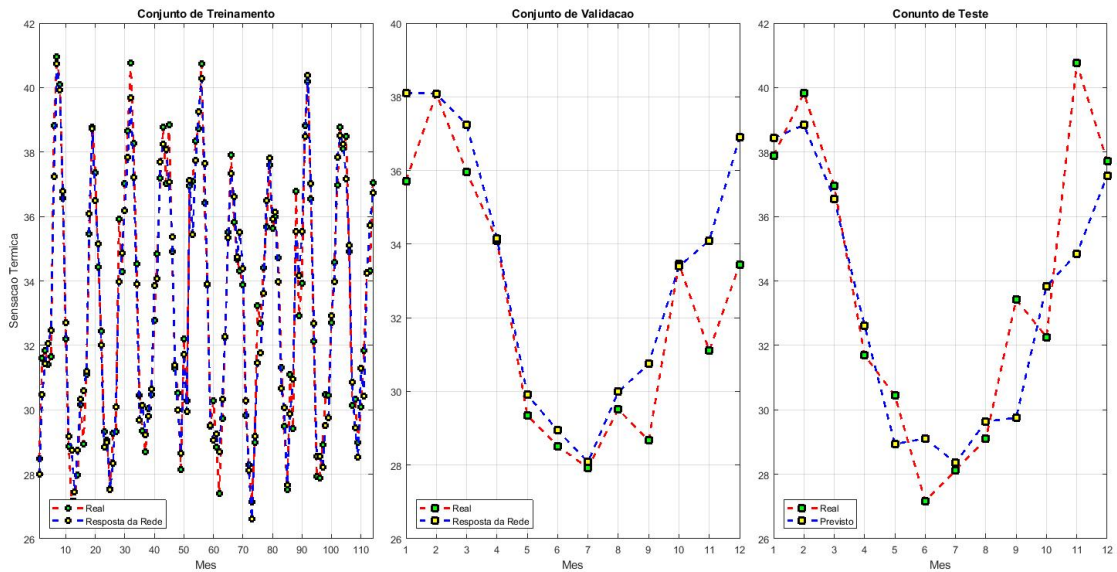
Figura 19: Quadro geral de testes. Da primeira para a última tríades de gráficos estão apresentadas na ordem crescente de experimentos que vão do 1 – 3.



(a) Experimento 4

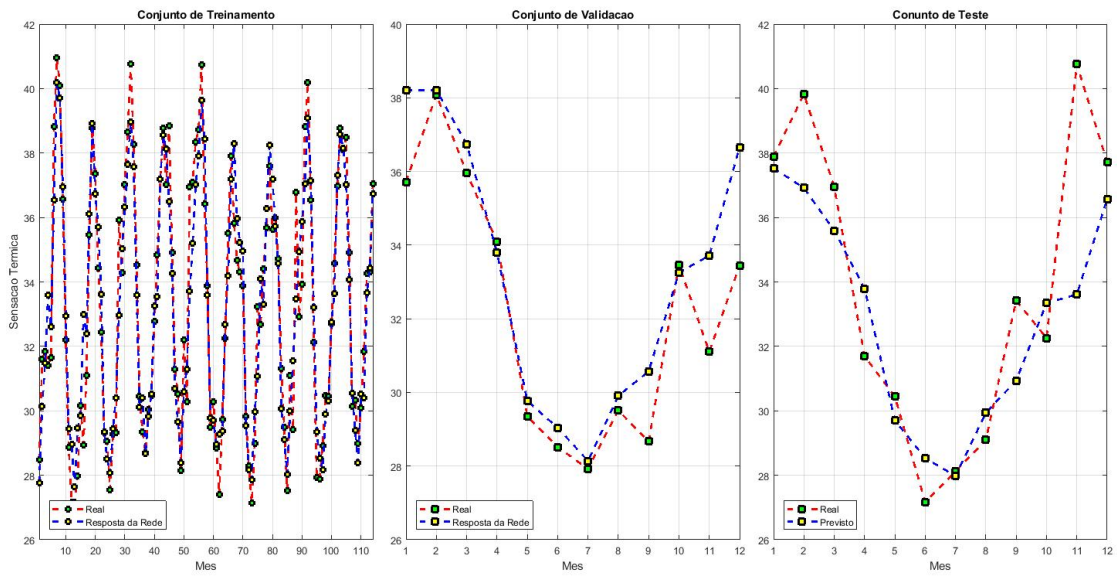


(b) Experimento 5

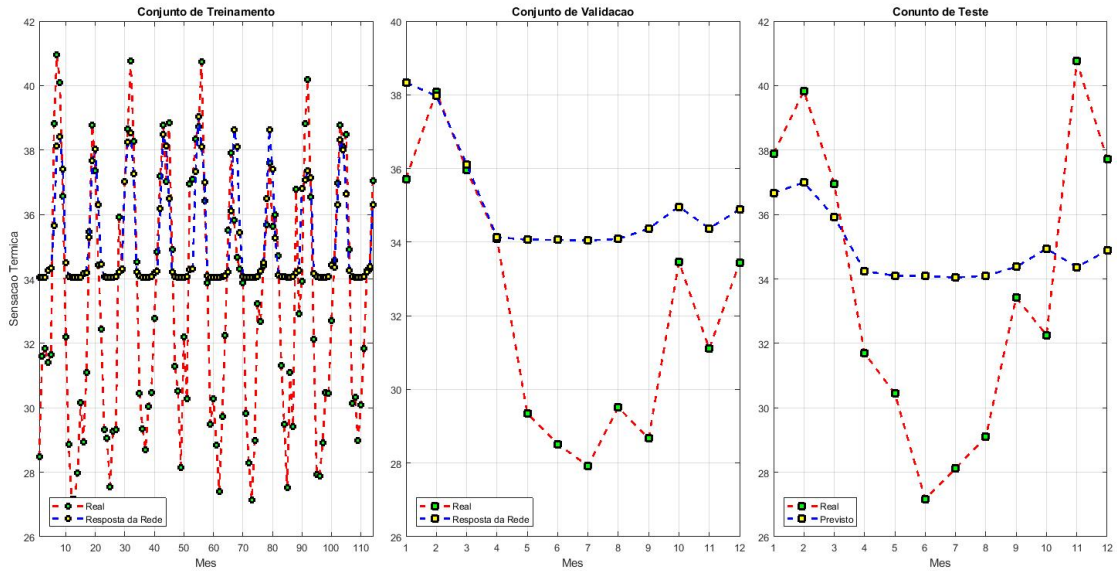


(c) Experimento 6

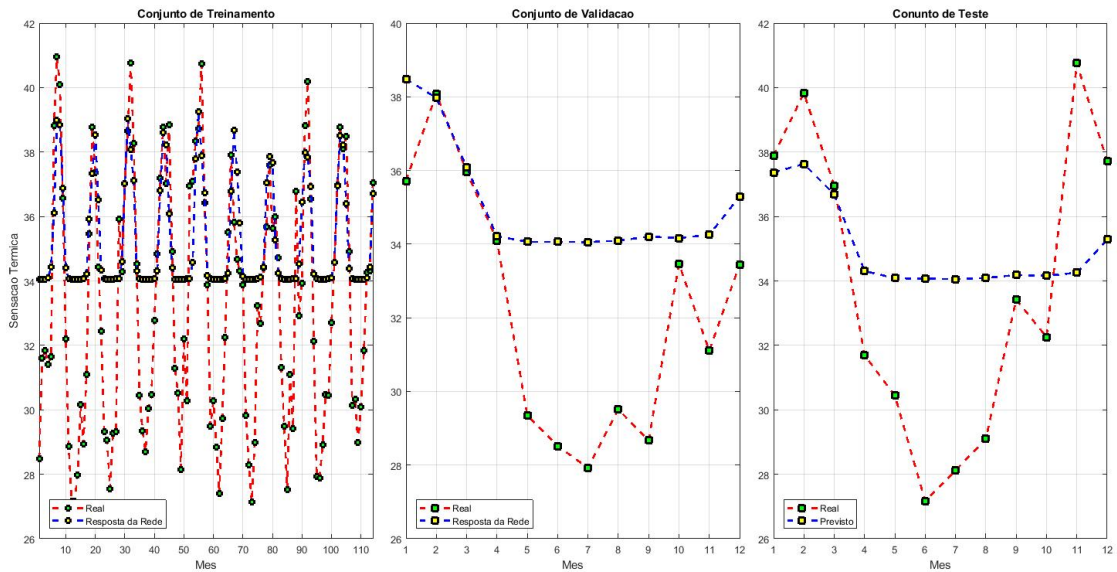
Figura 20: Quadro geral de testes. Da primeira para a última tríades de gráficos estão apresentadas na ordem crescente de experimentos que vão do 4 – 6.



(a) Experimento 7



(b) Experimento 8



(c) Experimento 9

Figura 21: Quadro geral de testes. Da primeira para a última tríades de gráficos estão apresentadas na ordem crescente de experimentos que vão do 7 – 9.