

UTILIZAÇÃO DO MÉTODO DE OTIMIZAÇÃO POR NUVEM DE PARTÍCULAS (PSO) NA ESCOLHA DA TOPOLOGIA DE REDES NEURAIS ARTIFICIAIS EMPREGADAS NA ESTIMAÇÃO DA VELOCIDADE DO VENTO

ERICK C. BEZERRA, ARTHUR P. DE S. BRAGA E RUTH P.S. LEÃO.

DEPARTAMENTO DE ENGENHARIA ELÉTRICA, UNIVERSIDADE FEDERAL DO CEARÁ
CAIXA POSTAL 6001 – CAMPUS DO PICI – CEP:60455-760
ERICK.BEZERRA@GMAIL.COM, ARTHURP@DEE.UFC.BR, RLEAO@DEE.UFC.BR

Resumo – Este trabalho utiliza o método de Otimização por Nuvens de Partículas (Particle Swarm Optimization – PSO) para estruturar Redes Neurais Artificiais (RNAs) que estimam a velocidade do vento em parques eólicos. As RNAs empregadas são do tipo NARX e FTDNN. Os resultados obtidos com estas RNAs são comparados com os obtidos com o modelo de Persistência, muito utilizado na prática. Foi utilizada uma série histórica com 45.658 dados de medição da velocidade do vento, em que 80% das medições foram selecionadas para a fase de treinamento e 20% para a validação. Como critérios de avaliação do desempenho das redes foram considerados as seguintes medidas de erro: MAE, RMSE e MAPE. Os resultados das previsões entre 1h e 6h são muito similares para todos os modelos, entretanto nas previsões para 12h e 48h existem diferenças entre os valores do MAPE que apontam a rede NARX como de maior aderência na projeção dos dados considerados.

Palavras-chave – Previsão de velocidade de vento, redes neurais artificiais, otimização por nuvens de partículas.

1 Introdução

Muitos países têm devotado elevada atenção no uso de fontes renováveis para atender a demanda crescente de energia e assegurar um desenvolvimento sustentável. Nas últimas duas décadas, a energia eólica surgiu como a principal tecnologia de energia verde. As projeções do *Global Wind Energy Council* (GWEC) indicam que até 2015 a capacidade mundial de potência instalada a partir da energia eólica deverá chegar a 449 GW, o dobro da potência instalada em 2010 [1]. A geração eolielétrica tem como característica a variabilidade da produção e o controle restrito da geração. A crescente penetração da energia eólica na matriz energética de muitos países, inclusive no Brasil, cria novas demandas para o setor elétrico nas áreas de planejamento, operação e mercado de energia. A previsão da geração eolielétrica é hoje uma necessidade para a operação confiável, segura e econômica dos sistemas de potência. A previsão pode ser usada para a combinação ótima da geração eólica e hidráulica. Pode ainda servir como base para quantificar a necessidade de reserva de potência para compensar a eventual redução de geração eólica. Gerenciar a variabilidade da geração eólica é o aspecto chave associado à integração ótima desta energia renovável na rede elétrica [2,3,4].

Definir o objetivo da previsão (Figura 1) é o primeiro passo para desenvolver um sistema de previsão. A previsão da geração eólica pode ser considerada em diferentes escalas de tempo dependendo da aplicação [9]. Uma previsão de milissegundos a poucos minutos, considerada como de curtíssimo prazo, é usada no controle de potência ativa do aerogerador. Para o planejamento da operação e mercado de energia, a previsão pode compreender uma faixa de 1 a 72 horas. Neste caso, a previsão auxilia na programação e despacho de unidades de geração convencionais. Esta escala de previsão é definida como de curto prazo. Para períodos de 5 a 7 dias a previsão pode ser considerada para o planejamento da manutenção das plantas eólicas, e é dita ser de longo prazo. A previsão com horizontes acima de 12 meses é usada no planejamento de recursos energéticos e em leilões de energia [21].

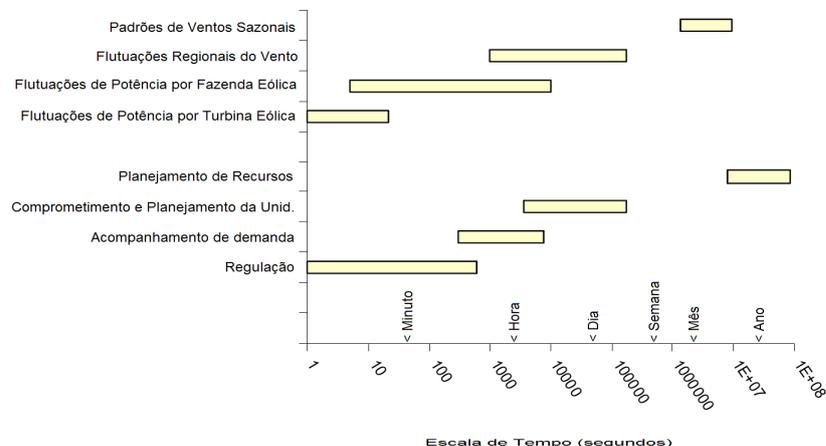


Figura 1 - Horizontes de tempo segundo a aplicação da previsão de geração eólica [21].

Para este trabalho, a previsão é utilizada para a operação de um aerogerador, com horizontes de tempo de: 1, 6, 12, 24 e 48 horas, e a frequência de atualização da previsão é horária. O objetivo do artigo é investigar o erro de previsão para diferentes métodos estatísticos que utilizem apenas dados históricos de velocidade do vento e apresentar o uso do método de Otimização por Nuvens de Partículas, do inglês PSO, como ferramenta para definir a melhor estrutura de uma Rede Neural Artificial (RNA). Assim, o artigo segue a seguinte estrutura: A Seção 2 apresenta o Modelo de Persistência para previsão dos ventos, que serve de referência para a comparação do desempenho das RNAs utilizadas, que são descritas na Seção 3; os dados de velocidade do vento e os indicadores de desempenho da previsão são tratados na Seção 4, e empregados na Seção 5 para avaliar os resultados obtidos; a Seção 6 encerra o artigo com as conclusões.

2 Modelo de Persistência

O modelo de persistência é considerado bastante eficaz na previsão de curta duração de velocidade do vento [16], e tem sido usado como referência para comparação de modelos de previsão. O modelo de persistência é definido como:

$$y(t) = y(t - i) \quad (1)$$

em que $y(t)$ é o valor previsto do vento em um instante t e o termo $y(t-i)$ é o valor da velocidade do vento no instante $t-i$.

O modelo de persistência em geral apresenta bons resultados, pois o tempo necessário para que mudanças atmosféricas ocorram é da ordem de várias horas. Para previsão de velocidades de vento em até 6 horas, o modelo de persistência é considerado como a referência a ser superada [23].

Algumas publicações comparam modelos de Redes Neurais Artificiais (RNAs) com o modelo de Persistência, e foi encontrado que normalmente as RNA se sobressaem em relação ao modelo de Persistência, mas que algumas vezes essa vantagem não é muito significativa [2,4,8].

3 Redes Neurais Artificiais

Muitos esforços têm sido devotados para desenvolver bons métodos de previsão do vento. As abordagens encontradas na literatura incluem métodos físicos, métodos estatísticos, modelos híbridos (físicos/estatísticos), e métodos baseados em técnicas de inteligência artificial [2,8,10,17]. Nos modelos baseados em RNA, as redes neurais do tipo MLP (*Multi-Layer Perceptrons*) [5,15], RBF (*Radial Basis Function*) [6] e Redes Neurais Recorrentes [12,18] têm sido usadas para esse fim.

3.1. RNA utilizadas na previsão da velocidade do vento

A RNA treinada para séries temporais tem a habilidade de modelar funções lineares e não lineares, com aplicação em campos diversos, incluindo detecção de transientes, reconhecimento de padrões, interpolação de funções, e previsão de séries temporais.

A RNA é uma tecnologia promissora na previsão da velocidade do vento. Lapedes e Farber (1987) propuseram um modelo de RNA do tipo *feed-forward* utilizando o algoritmo *back-propagation* na previsão da velocidade do vento. Song (2000) desenvolveu uma RNA que utiliza atrasadores e que a soma dos pesos das entradas deve ser igual a 1. O modelo apresentou bom resultado para previsões da hora seguinte quando os dados do vento não oscilam abruptamente. Outras técnicas são aplicadas para melhorar o resultado das redes neurais, Li, Shi e Zhou (2010) combinam o resultado de várias redes para realizar a previsão.

A rede MLP (Multilayer Perceptron) com atrasadores na entrada é conhecida pela sigla FTDNN (Focused Time Delay Neural Network) (Figura 2). A rede FTDNN é uma rede *feed-forward* multicamadas cujos pesos sinápticos são ajustados de acordo com o algoritmo de retro-propagação convencional [20].

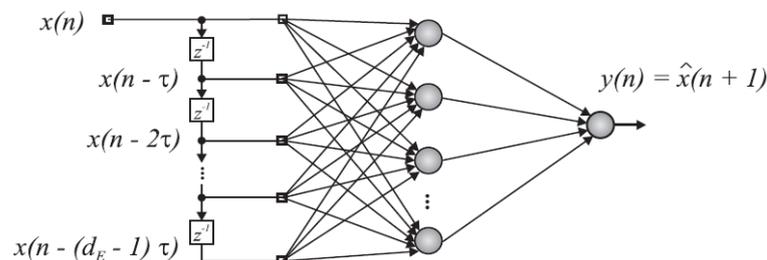


Figura 2 - Rede FTDNN com arquitetura genérica [11].

$x(n)$ é um vetor que contém d_E elementos da série contados a partir do elemento atual $x(n)$, espaçados um do outro de τ unidades de tempo, y é saída do neurônio e \hat{x} é a variável escalar predita, z^{-1} é apenas uma indicação de atraso no tempo.

As RNA recorrentes, como por exemplo a rede NARX [7,19], são redes que contêm conexões sinápticas realimentadas da saída da rede (Figura 3).

Durante os treinamentos foram modificadas as maneiras com que os atrasadores eram apresentados às redes, não sendo observado mudanças significativas no resultado deste trabalho.

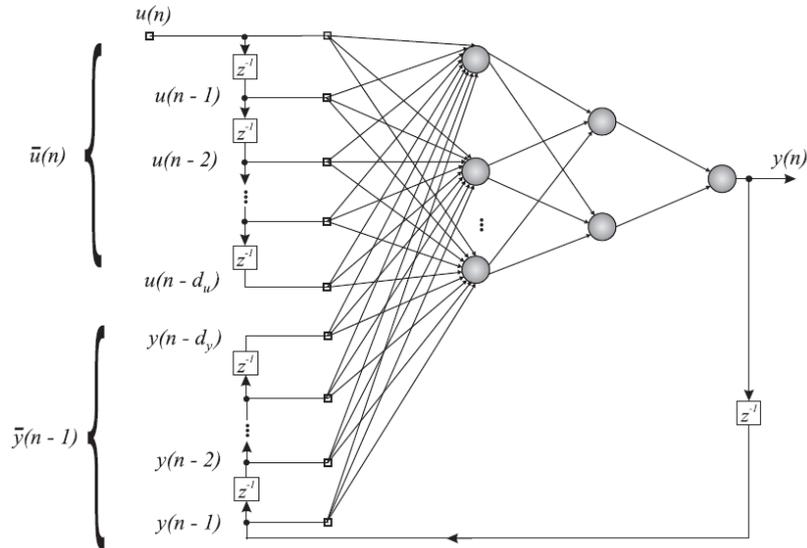


Figura 3 - Rede NARX com arquitetura genérica [11].

O regressor de entrada $u(n)$ contenha d_u amostras da variável observada $x(n)$, espaçadas de $\tau > 0$ unidades de tempo, enquanto que o regressor de saída $y(n-1)$ contém valores reais ou estimativas da mesma variável, porém amostradas em instantes consecutivos.

3.2. Algoritmo de aprendizagem Backpropagation

Todas as redes usadas no modelo de previsão de velocidade do vento foram treinadas utilizando o *back-propagation* ou retro-propagação [22].

Os pesos iniciais para as unidades de entrada foram escolhidos aleatoriamente. Contudo, para as camadas de saída e intermediária os pesos foram atualizados conforme (2) e (3), respectivamente.

$$m_{ki}(n+1) = m_{ki}(n) + \eta \delta_k(n) y_i(n) \quad (2)$$

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_i(n) x_j(n) \quad (3)$$

em que η é a taxa de aprendizado e m e w são os pesos.

O cálculo dos gradientes locais foi realizado com base em (4).

$$\delta(n) = e(n) \varphi' [u(n)] \quad (4)$$

em que $e(n)$ é o erro entre a saída desejada e a resposta gerada pela rede, enquanto $\varphi'[u(n)]$ definida em (5) representa a derivada da função de ativação (6), neste trabalho, a função sigmóide.

$$\varphi' [u(n)] = \varphi [u(n)] * (1 - \varphi [u(n)]) \quad (5)$$

$$\varphi [u(n)] = \frac{1}{(1 + e^{-\lambda(u(n))})} \quad (6)$$

Para o cálculo dos gradientes locais das camadas escondidas, utilizou-se (7).

$$\delta_i(n) = \varphi' [u_i(n)] \sum m_{ki} \delta_k(n) \quad (7)$$

3.3. Otimização por Nuvem de Partículas (Particle Swarm Optimization – PSO)

O PSO é uma técnica de inteligência de enxames baseada na tentativa de modelar um bando de pássaros procurando comida [25 e 26]. Foi desenvolvido por Kenedy e Eberhart em 1995. O algoritmo básico do PSO, denominado modelo canônico do PSO, pode ser descrito em notação vetorial como a seguir:

$$v_{k+1} = a \otimes v_k + b_1 \otimes r_1 \otimes (p_1 - x_k) + b_2 \otimes r_2 \otimes (p_2 - x_k) \quad (8)$$

$$x_{k+1} = c \otimes x_k + d \otimes v_{k+1} \quad (9)$$

sendo \otimes é o produto vetorial elemento por elemento; v é o vetor velocidade da partícula; x é o vetor posição da partícula; a é o fator de momento (peso de inércia); p_1 é a melhor posição anterior da partícula (P_{best}); p_2 melhor posição anterior do enxame (G_{best}); b_1 , b_2 são respectivamente a constante cognitivo e social; r_1 , r_2 são vetores de números randômicos no intervalo $[0,1]$ e c , d são vetores coeficientes.

As partículas são inicializadas com posições aleatórias e velocidades nulas, calcula-se o valor de *fitness* ou de aptidão (função a ser otimizada), compara-se os *fitness* com os melhores valores do indivíduo (P_{best}) e do enxame (G_{best}), velocidades e posições são atualizadas e repetido o ciclo até o critério de parada ser atingido.

4 Dados de velocidade do vento e indicadores de desempenho da precisão da previsão

4.1 Dados de velocidade do vento

Os dados de vento foram coletados do site do *Alternative Energy Institute (AEI)*, da *West Texas A&M University*. Foram priorizados medições de velocidade a 100 m de altura. Os dados são relativos aos anos de 2003 à 2008. O site escolhido foi o *Tall Tower South – Site #51 (Sweetwater)*, latitude $32^\circ 24' 42''$ e longitude $100^\circ 21' 16''$. Também foi verificada a inexistência de dados inconsistentes (velocidades negativas, ou valores fora da curva, como ventos acima de 30 m/s) e identificada a quantidade de medições não realizadas dentro do período coletado, que totalizou 2,34% das medições, não sendo utilizado então nenhum método de preenchimento de falhas.

4.2 Indicadores de desempenho da previsão

Vários indicadores de performance de modelos foram propostos e empregados na avaliação da previsão de velocidade do vento, mas nenhum indicador é reconhecido como um indicador universal padrão. Os erros calculados neste estudo são: *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, e o *Mean Absolute Percentage Error (MAPE)*.

$$MAE_j = \frac{1}{T} \sum_{t=1}^T |y_t - f_{j,t}| \quad (10)$$

$$RMSE_j = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - f_{j,t})^2} \quad (11)$$

$$MAPE_j = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - f_{j,t}}{y_t} \right| \quad (12)$$

sendo: y_t e $f_{j,t}$ denotam as observações e o valor previsto pelo modelo j , e T é o número de dados usado para avaliação e comparação.

O MAE (Equação 10) mede a média absoluta do erro em um conjunto de previsões. O MAE é uma medição linear, todas as diferenças estão recebendo o mesmo peso igualmente na média. O RMSE (Equação 11) é uma medição quadrática do erro quadrático médio. Como os erros são elevados ao quadrado antes de ser calculada a média, o RMSE fornece um peso maior aos erros maiores. Isso significa que o RMSE é mais útil quando grandes erros são particularmente indesejáveis. O MAPE (Equação 12) é a diferença entre o valor atual e o previsto dividido pelo valor atual. O módulo deste cálculo é então somado para cada previsão e dividido pelo número total de pontos previstos.

As estatísticas MAE e RMSE dependem da escala da variável dependente é invariante à escala da variável dependente. Quanto menor as estatísticas, melhor é a habilidade do modelo de fazer previsões.

5 Avaliação dos resultados

Utilizou-se dados correspondentes a 36.242 pontos horários para treinar as Redes Neurais FTDNN e NARX que foram implementadas utilizando o MATLAB™. Por tentativa e erro foi encontrada a estrutura, utilizada em todos os horizontes de

previsão, de 3 neurônios na única camada intermediária, 7 entradas e 7 atrasos. Uma série de 9.416 pontos horários, i.e., cerca de 26% das medições, foram usados para a validação dos modelos RNA utilizados. A comparação dos erros dos diferentes métodos de previsão implementados são apresentados em gráfico e tabelas como se seguem (Tabelas 1 a 5).

Tabela 1 - Erros calculados para previsão de 1 hora.

	1 hora MAE	RMSE	MAPE
FTDNN	0,0415	0,0545	0,1779
NARX	0,0360	0,0489	0,1616
PERSIST	0,0299	0,0412	0,1461

Tabela 2 - Erros calculados para previsão de 6 horas.

	6 hora MAE	RMSE	MAPE
FTDNN	0,1865	0,2322	0,5598
NARX	0,0893	0,1148	0,3937
PERSIST	0,0839	0,1067	0,4323

Tabela 3 - Erros calculados para previsão de 12 horas.

	12 hora MAE	RMSE	MAPE
FTDNN	0,2128	0,2649	0,6234
NARX	0,1119	0,1431	0,4545
PERSIST	0,1074	0,1345	0,5680

Tabela 4 - Erros calculados para previsão de 24 horas.

	24 hora MAE	RMSE	MAPE
FTDNN	0,2357	0,2875	0,6693
NARX	0,1193	0,1523	0,4698
PERSIST	0,1053	0,1360	0,5476

Tabela 5 - Erros calculados para previsão de 48 horas.

	48 hora MAE	RMSE	MAPE
FTDNN	0,2288	0,2805	0,6470
NARX	0,1198	0,1532	0,4732
PERSIST	0,1175	0,1486	0,6158

As Figuras 9 a 12 apresentam para cada horizonte de previsão e método avaliado a aderência entre valores medidos e projetados. Foi selecionado um intervalo aleatório de amostras entre 5500 e 5600 para criar as figuras. Em azul estão as medições e em vermelho são previsões. Nas Figuras 9 e 10 pode ser observado como o modelo de Persistência perde sua “capacidade” de aderência para um horizonte de previsão maior que 6h.

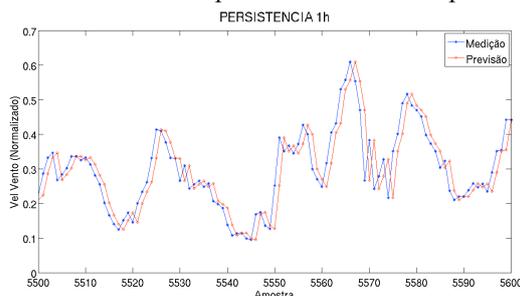


Figura 9 - Modelo de Persistência para previsão da hora seguinte.

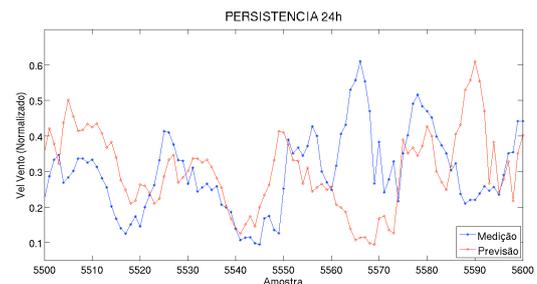


Figura 10 - Modelo de Persistência para previsão da 24ª hora seguinte.

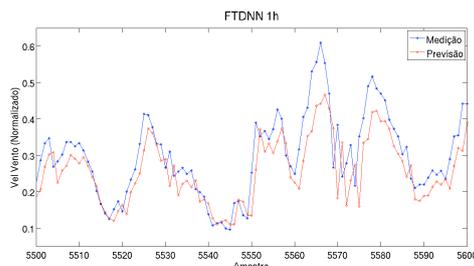


Figura 11 - Rede FTDNN com 7 atrasadores em série para previsão da hora seguinte.

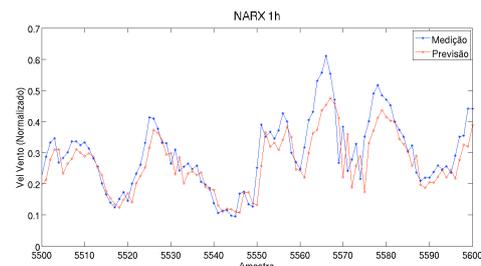


Figura 12 - Rede NARX com 7 atrasadores em série para previsão da hora seguinte.

Como citado anteriormente, a configuração das redes neurais foi obtida por tentativa e erro. Para definir uma estrutura ótima para as redes, foi usado o método PSO.

Uma das primeiras aplicações do PSO foi no treinamento de redes feedforward [25,26]. Desde então, tornou-se um algoritmo popular para o treinamento de redes neurais [27-29], particularmente no uso como um otimizador global [30,31].

A idéia inicial era utilizar o PSO tanto para estruturar a rede, definindo número de neurônios e entradas da rede, como treiná-la. No entanto, o tempo consumido durante o treinamento foi bem maior que o do Backpropagation, então decidiu-se por usar o método clássico do PSO para definir o número de neurônios da camada intermediária e o número de entradas e o método de retro-propagação para treinar a rede, e como função de fitting foi usado o MSE do conjunto de treinamento. Como a rede que apresentou melhor resultado anteriormente foi a NARX, o trabalho realizado com o PSO foi feito com esta rede.

Para cada faixa de previsão utilizou-se o mesmo número de épocas (10), gerações (10), partículas (15) e velocidades iniciais (nulas). Para todas as redes considerou-se um intervalo para o número de neurônios na camada intermediária (entre 1 e 19) e o número de entradas (entre 1 e 19), baseado já na experiência obtida na tentativa e erro, evitando assim a *over fitting*. Como os pesos iniciais são aleatórios foram utilizados 3 conjuntos diferentes de pesos para cada partícula.

As tabelas 6 a 10 apresentam alguns dos resultados obtidos com a rede NARX e estrutura de nuvem de partículas como descrita acima. Na coluna da esquerda tem-se a identificação dos arquivos, onde a primeira linha mostra a estrutura utilizada na tentativa e erro (T.E.) Para medir a aderência ou erro do conjunto de dados de validação foi usada a medida MSE designada como MSEV. A coluna NCI indica o número de neurônios da camada interna e Ninput o número de entradas da rede. Os resultados obtidos apresentaram várias estruturas, com resultados distantes do obtido através do método por tentativa e erro.

Tabela 6 – Estruturas (PSO) para previsão de 1 hora.

1h	MSEV	NCI	Ninput
T.E.	0.0024	3	15
05.09.03	0.0024	3	3
05.09.02	0.0028	11	3
05.09.01	0.0381	9	7

Tabela 7 – Estruturas (PSO) para previsão de 6 horas.

6h	MSEV	NCI	Ninput
T.E.	0.0131	3	15
01.09.01	0.0596	5	29
01.09.02	0.0322	9	17
02.09.01	0.0369	16	33

Tabela 8 – Estruturas (PSO) para previsão de 12 horas.

12h	MSEV	NCI	Ninput
T.E.	0.0204	3	15
30.08.01	0.5587	15	3
31.08.01	0.0820	9	3
31.08.02	0.5576	20	3

Tabela 9 – Estruturas (PSO) para previsão de 24 horas.

24h	MSEV	NCI	Ninput
T.E.	0.0231	3	15
28.08.01	0.0266	19	3
29.08.01	0.0323	19	37
29.08.02	0.0712	12	3

Tabela 10 – Estruturas (PSO) para previsão de 48 horas.

48h	MSEV	NCI	Ninput
T.E.	0.0234	3	15
25.08.01	0.0264	9	35
26.08.01	0.0821	8	3
28.08.01	0.0267	15	3

Quando para uma das partículas foi fixado os números da etapa da tentativa e erro (neurônios da camada interna, número de entradas e conjunto de pesos) a estrutura das redes sofreu alteração, sendo o resultado apresentado na Tabela 11 somente para a rede NARX para 1h.

Tabela 11 – Estruturas (PSO) para previsão de 1 hora com o uso de uma partícula como “referência”.

1h	MSEV	NCI	Ninput
T.E.	0.0024	3	15
07.09.01	0.0063	3	5
08.09.01	0.0053	3	9
08.09.02	0.0234	3	5

Com a inclusão da partícula de “referência” os resultados do PSO tenderam a se concentrar num “espaço” solução.

6 Conclusão

Pode-se concluir que utilizando somente os dados de velocidade de vento de um dado sítio, os modelos de previsão de velocidade do vento como Persistência, FTDNN e NARX apresentaram resultados próximos em todos os horizontes de tempo.

Para previsões iguais ou maiores que 6h existe uma crescente diferença no valor de erro calculado pelo método MAPE, todos favorecendo o modelo NARX, chegando a diferença de 14,26% entre os resultados desta rede e o modelo de Persistência na previsão para 48h. Para todos os horizontes o modelo NARX obteve melhor resultado que o FTDNN.

A pesquisa realizada com o PSO apresentou resultados similares aos obtidos por tentativa e erro.

7 Agradecimentos

Os autores agradecem a FUNCAP pelo apoio financeiro.

8 Referências

- [1] Global Wind Energy Council (2010). Global Wind Report 2010. Annual Market Update 2010, pp. 18-21.
- [2] Lei, M., Shiyang, L., Chuanwen, J., Hongling, L. and Yan, Z (2009). A review on the forecasting of wind speed and generated power. *Renew Sustain Energy Rev*, 13(4): 915–20.
- [3] Luickx, P.J., Delarue, E.D. and D'haeseleer, W.D (2008). Considerations on the backup of wind power: operational backup. *Applied Energy*, 85(9): 787–99.
- [4] Sfetsos, A (2000). A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renew Energy*, 21(1):23–35.
- [5] Alexiadis, M (1998). Short-term forecasting of wind speed and related electrical power. *Solar Energy*, 63(1):61–8.
- [6] Beyer, H.G., Degner, T., Haussmann, J., Homan, M. and Rujan, P (1994). Short term forecast of wind speed and power output of a wind turbine with neural networks. In: *Proceeding on the 2nd European Congress on Intelligent Techniques and Software Computing*. Aachen, Germany.
- [7] Chen, S., Billings, S.A. and Grant, P.M (1990). Non-linear system identification using neural networks. *International Journal of Control*, 51(6):1191-1214.
- [8] Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H. and Feitosa, E (2008). A review on the young history of the wind power short-term prediction. *Renew Sustain Energy Rev*, 12(6):1725–44.
- [9] Cutler, N (2006). *Wind Energy Forecasting Issues Paper*. Version 12.
- [10] Fadare, D.A (2010). The application of artificial neural networks to mapping of Wind speed profile for energy application in Nigeria. *Appl Energy*, 87(3): 934–42.
- [11] Menezes Júnior, J.M (2006). *Redes Neurais Dinâmicas para Predição e Modelagem Não-Linear de Séries Temporais*. Dissertação (Mestrado em Engenharia de Teleinformática) - Universidade Federal do Ceará.
- [12] Kariniotakis, G., Stavrakakis, G.S. and Nogaret, E.F (1996). Wind power forecasting using advanced neural network models. *IEEE Trans Energy Convers*, 11(4): 762–7.
- [13] Lapedes, A. and Farber, R (1987). *Nonlinear signal processing using neural networks: forecast and system modeling*. Los Alamos National Lab. Los Alamos, NM. Technical report LAUR872662.
- [14] Li, G., Shi, J. and Zhou J (2010). Bayesian adaptive combination of short-term wind speed forecasts from neural networks models. *Renewable Energy*, 36(1): 352-9.
- [15] Lin, L., Eriksson, J.T., Vihriala, H. and Soderlund, L (1996). Predicting wind behavior with neural networks. In: *Proceeding of the European Wind Energy Conference 1996*. Goteborg, Sweden, p. 655–8.
- [16] Makarov, Y., Hawkins, D., Leuze, E. and Vidov, J (2003). California ISO Wind Generation Forecasting Service Design and Experience. *Proceedings of American Wind Energy Association Conference*, Oregon, USA.
- [17] Morales, J.M., Mínguez, R. and Conejo, A.J (2010). A methodology to generate statistically dependent wind speed scenarios. *Appl Energy*, 87(3):843–55.
- [18] More, A. and Deo, M.C (2003). Forecasting wind with neural networks. *Mar Struct*, 16(1):35–49.
- [19] Narendra, K.S. and Parthasarathy, K (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*. 1:4-27.
- [20] Principe, J.C., Euliano, N.R., Lefebvre, W.C (2000). *Neural Adaptive Systems: Fundamentals Through Simulations*. John Wiley and Sons.
- [21] Piwko, R., Bai, X., Clark, K., Jordan, G., Miller, N. and Zimmerlin, J (2005). *The Effects of Integrating Wind Power on Transmission System Planning, Reliability, and Operations*. Report on Phase 2: System Performance Evaluation, Albany, New York, USA, pp. 16.
- [22] Rumelhart, D.E., Hinton, G.E. and Williams, R.J (1986). Learning representations by back-propagating errors. *Nature* 323: 533-536.
- [23] Siebert, N (2008). *Development of Methods for Regional Wind Power Forecasting*. pp. 37-38.
- [24] Song, Y.D (2000). A new approach for wind speed prediction. *Wind Engineering*, Vol. 24, No.1, pp.35–47.
- [25] Eberhart, R. and Kennedy, J (1995). A new optimizer using particle swarm theory. In *Micro Machine and Human Science*. MHS '95, Proceedings of the Sixth International Symposium on, pages 39-43.
- [26] Kennedy, J (1997). The particle swarrn: social adaptation of knowledge. In *Evolutionary Computation*. IEEE International Conference on, pages 303-308.
- [27] Corne, D., Dorigo, M. and Glover, F (1999). *New Ideas in Optimization*. McGraw-Hill.

- [28] Eberhart, R., Simpson, P.K., and Dobbins, R.W (1996). Computational Intelligence PC Tools. Academic Press Professional.
- [29] Bergh, F. van den. and Engelbrecht, A. P (1999). Particle swarm weight initialization in multi-layer perceptron artificial neural neural networks. In Artificial Intelligence. Proceedings of the International Conference on, pages 42–45.
- [30] Ismail, A. and Engelbrecht, A.P (2000). Global optimization algorithms for training product unit neural networks. In Neural Networks. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on, volume 1, pages 132–137 vol.1.
- [31] Bergh, F. van den. and Engelbrecht, A. P (2001). Training product unit networks using cooperative particle swarm optimisers. In Neural Networks. Proceedings. IJCNN '01. International Joint Conference on, volume 1, pages 126–131.
- [32] Kennedy, J. and Eberhart, R (1995). “Particle swarm optimization,” in Proc. Of the IEEE International Joint Conference on Neural Networks. IEEE Press, pp. 1942–1948.