Proceedings of International Joint Conference on Neural Networks, Montreal, Canada, July 31 - August 4, 2005

Artificial Neural Networks for Temporal Processing Applied to Prediction of Electric Energy Generation in Small Hydroelectric Power Stations

Paulo Cesar Endo Joaquim
Centro de Ciências Exatas, Ambientais e de Tecnologias
PUC-Campinas
Campinas, SP, Brazil
E-mail: pcejoaquim@yahoo.com.br

João Luís Garcia Rosa
Centro de Ciências Exatas, Ambientais e de Tecnologias
PUC-Campinas
Campinas, SP, Brazil
E-mail: joaoluis@puc-campinas.edu.br

***Abstract*— The purpose of this work is to present a computational prediction of temporal series through Artificial Neural Networks (ANN) with temporal features based on short-term memory structures and episodic long-term memory. The connectionist prediction is applied to a Brazilian small hydroelectric power station, with generation capacity of 15 MWh, because conventional prediction statistical techniques show inadequacy in relation to noise, acquisition fails, and need for generalization, when applied to this model. Departing from the proposed system, it is intended also to develop, in the future, a non-linear complex system, employing ANNs, with the inclusion of new variables in the decision process, in addition to the episodic memory model, which is considered computationally feasible with the current available resources.**

I. Introduction

There are several applications of artificial neural networks to electric systems, because they are an excellent alternative to solve control problems, prediction, maintenance, and optimization in non-linear complex electric systems [3].

In electric systems, one of the main applications of ANNs is the load/demand prediction, because the key objective of an electric energy enterprise is to offer a high quality service to its consumers, assuring a reasonable charge, consequence of a better distribution electric management. So, it should allocate, in an optimized way, the operational and investment resources in the distribution network. A correct knowledge of the load profile is of vital importance to this optimization, because prediction fails can cause an increasing in operational costs. The demand is directly related to the electric energy consumption, consequently to variables that influence directly this amount of used energy [5].

Prediction that involves electric energy generation depends on water resources of hydroelectric power stations, and on reservoir water availability; in addition, electric energy generation in small stations is not controlled by the Brazilian government, unlike large hydroelectric power stations [1].

This is complicated because it adds a random and temporal factor to the generation capacity, since the total water availability in the station is not guaranteed and, the possibility of generating energy at any instant through total power is not guaranteed either.

The neural network technique application is considered a very interesting way of modeling generation prediction process, because it is allowed to "learn" with the historic generation the probable behavior in the future generation. In the case studied, there is a historic series of daily generation, which ranges from the year 2000 to the year 2004, period considered relatively short, regarding annual seasonality. This period can entail great uncertainty in a purely statistical prediction criterion.

Neural networks applied to temporal series prediction present a kind of short-term memory. This memory makes possible the networks treat input and output patterns that vary in a specific time period.

II. Memory

The word *memory* is originated from Latin *memoria*, from *memor*, which means 'remembering', that is, the faculty of retaining or reacquiring ideas, images, expressions, and knowledge acquired previously, reporting to memories and reminiscences [6]. Memory is the place where learned knowledge is stored [11]. Several definitions of memories have been proposed in many terms such as data memory, recognizing memory, short-term memory, work memory, long-term memory, processing memory, motor memory, contextual memory, etc. Table I displays different types of memory and the division regarding the retention time. According to [6], explicit and implicit memories are considered long-term memories, that store knowledge, while operational memory is short-term, because the information is stored temporarily and useful only to immediate reasoning and problem resolution.

The information processing is serial, that is, it is necessary that information passes first through the sensorial storing, then through short-term memory, and finally through long-term memory. The passing of information from short-term memory to long-term memory depends on some control processes, like information repetition and adequate information encoding [8].

TABLE I
TYPES, SUBTYPES, AND MEMORY FEATURES. ADAPTED FROM [6]

Retention time	Types	Subtypes	Features
Long-Term	Explicit or Declarative	Episodic	Has temporal reference; sequenced facts memory
		Semantic	Involves non-temporal concepts; cultural memory
Long-Term	Implicit or non-declarative	Perceptual	Represents images with no known meaning; pre-conscious memory
		Procedural	Habits, abilities, and rules
		Associative	Associate two or more stimuli or a stimulus to a certain response
		Non-associative	Attenuates an answer or increases it through the same stimulus repetition
Short-Term	Operational		Allows reasoning and behavior planning

Considering the memory nature, there is basically two types of memory: the declarative or explicit memories, which can be reported and recognized as past-knowledge retention memory, and the implicit or non-declarative memories, also known as habits, that are acquired and evoked in a quasi-automatic way. Memories of a text, a fact, an event, and a face are explicit memories and memories of the way of walking or swimming, that is, motor coordination, are considered implicit memories [10].

Explicit memory encodes the information about events and facts, and its formation depends on cognitive processes such as evaluation, comparison, and inference [11]. This type of memory can be recovered by remembering acts and associates everything that can be evoked by means of words, which is the reason of the term declarative.

The explicit memory is divided into:

- *Episodic*: allows mentally backdating in time and recovering past events or experiences. An example of this type of memory is the recovering of past events, as the last birthday or a son's birth, allowing a re-experience in the event personal view involving several standpoint variables as smell, sound, taste, and feelings [12]. This ability to travel mentally through time makes a person revive experiences taking again situations occurred in the past and projecting them mentally anticipating the future through imagination. This bridge between past experiences and future predictions is one of the most important memory mechanisms.
- *Semantic*: refers to facts, personal information, or autobiographic retention that happen when non-temporal concepts are involved [6]. This type of memory is not dated, that is, the information is stored without the way and the time it is inserted in memory files. Refers to acquired knowledge about the world, like conducting ethical principles, playing chess, or knowing which is the capital of Brazil [4].

Implicit memory has an automatic and reflexive quality and its formation and remembering are not dependent on the capacity of having or being aware of cognitive processes. This type of memory is accumulated slowly after repetitions and can not be expressed by words, because it is connected to perceptual and motor abilities and learning of certain types of procedures and rules, as grammar [11]. This type of memory is also known as procedure memory [4].

There is also the operational memory, that unlike other previous types of memory is considered a short-term memory by retention time. In more recent approaches, the operational memory is also called primary or work memory, since it refers to temporary information storing during a period, sufficient to immediate reasoning and problem resolution. This memory can be extended for minutes or hours by memorization process, that is, by mental repetition of information content. This repetition easies also the information transference to the long-term storing system, for instance, the explicit memory [4].

III. UTILIZED TECHNIQUE

The construction of a non-linear dynamic system employing ANN is very common, and shows a clear responsibility separation: the static network is associated to the non-linearity and the memory is associated to time. This ANN utilization gives good results in temporal series prediction problem solutions, revealing, in many cases, a better performance than in purely statistical models [13].

According to Elman [7], "a neural network should have memory in order to be dynamic". In an ANN for temporal processing, the memory, depending on the retention time, can be divided in "short-term" memory and "long-term" memory. The long-term memory is inserted in an ANN by means of learning, in which information content of training data set is stored in network synaptic weights. If the task has a temporal dimension, it needs a short-term memory to make the ANN dynamic in relation to time [13].

The proposed system employs windowing techniques through multilayer perceptron FIR network formalizations. The windowing method representation, employing FIR network concepts, is given by the utilization of filters of lengths equal to desired window length, as can be seen in figure 1. According to Haykin [9], a MLP network whose hidden and output layer neurons are based on a finite-impulse duration response (FIR) is a FIR MLP.

This technique allows the incorporation of a short-term finite memory, in a convenient and controlled time period for

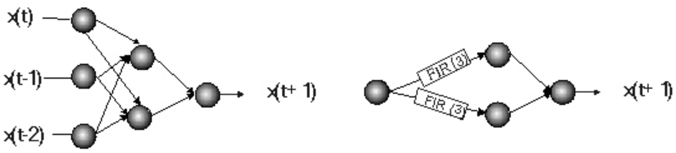

Fig. 1. Windowing method and FIR concepts equivalence

the energy generation prediction, being able of learning a non-stationary behavior. In this kind of network, the introduction of a short-term memory is controlled, giving recent and temporally previous input values to the neurons. This is a crucial factor for choosing FIR MLP instead of recurrent networks that use infinite time impulse filters, IIR - infinite impulse response, due to the employment of feedback loops by its architecture in order to create a short-term memory.

A FIR network with episodic memory is also employed, as shown in figure 2. The episodic memory allows the ANN associate the time in which the energy is generated, creating a temporal chronological link to the FIR network data and making available a bridge between the training past experience and the network future predictions.

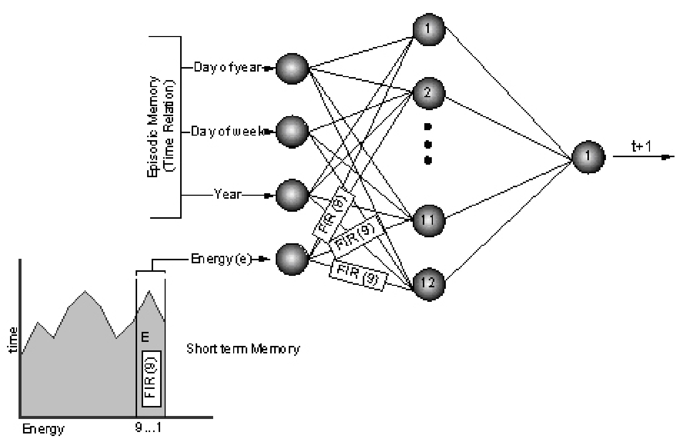

Fig. 2. Details of FIR network with episodic memory

The employed architecture is presented in figure 3. For the input and output layers one have to define the number of units per layer, adequate to the model to be developed. For the hidden layer, there is no rules about the number of layers and units [2]. The network is composed by the following structure:

- *Input layer*: the FIR network uses a 10-unit layer, which has enough number of neurons to provide to the ANN a short-term memory in the adequate prediction time period for the model. For the episodic memory implementation, three units represent the day of the week, the day of the year, and the year (see figure 6 for details).
- *Hidden layer*: only one 12-unit hidden layer with a number of units next to the input layer in order to keep small the training time. A large number of neurons could make the network memorize the training data (over fitting) instead of extracting the general features that allow generalization. A small number of neurons could make the network spend much time trying to find an optimal representation, and not converging [2].
- *Output layer*: only one unit corresponding to the prediction $x(t+1)$.

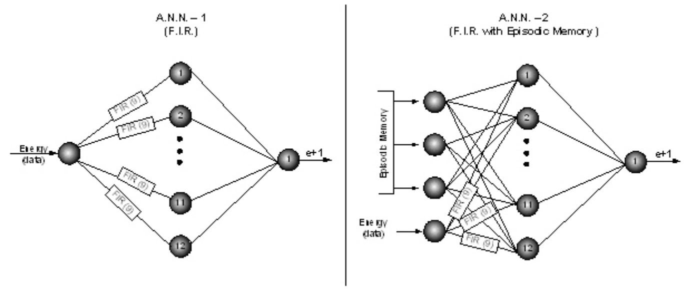

Fig. 3. An ANN architecture: FIR network on the left and FIR network with episodic memory on the right

IV. Training

The electric energy generation prediction is based on historic generation of a small station with capacity of 15 MWh generation. This small hydroelectric power station has a small water reservoir and depends, almost exclusively, on the rain volume of the water resource in installation location.

The data utilized in training, as well as the data for comparison of prediction with generation, can be visualized in figures 4 and 5, which show two diagrams:

- *Generated energy in MWh diagram*: history of the generated energy between January 2002 until November 2004 (figure 4);
- *November month diagram*: diagram shows the electric energy generation occurred in November, years 2002, 2003, and 2004; the November 2004 data are not used in training, since they are utilized to compare the predictions of two implemented ANN architectures (figure 5).

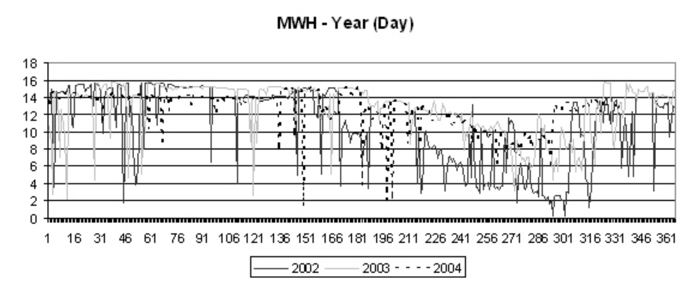

Fig. 4. Diagram with generation historic data of years 2002, 2003, and 2004 in daily period

The network structure, combined with energy generation data from the small hydroelectric power station, is shown in figure 6. This figure illustrates the network architecture, with or without episodic memory, regarding the generated energy data in the month of November, 2004. From this table it is possible to visualize how the temporal shift of FIR network

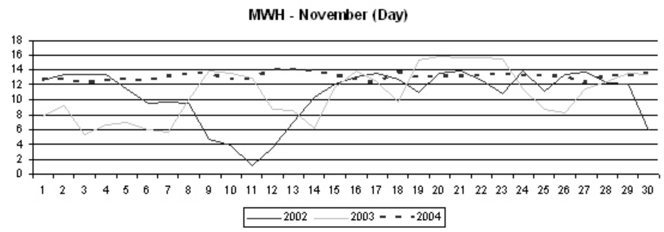

Fig. 5. November diagram

Input Layer														Output Layer	
Short-Term Memory											Episodic Memory				
Neuron Unit 1	Neuron Unit 2	Neuron Unit 3	Neuron Unit 4	Neuron Unit 5	Neuron Unit 6	Neuron Unit 7	Neuron Unit 8	Neuron Unit 9	Neuron Unit 10		Neuron Unit 11	Neuron Unit 12	Neuron Unit 13	Neuron Unit 1	
FIR (10)	FIR (10)	FIR (10)	FIR (10)	FIR (10)	FIR (10)	FIR (10)	FIR (10)	FIR (10)	FIR (10)	Representation of Day	Day of Week Value	Day of Year Value	Year Value	Prediction (t)	Representation of Day
12.8	12.7	12.3	12.5	12.7	12.6	13.1	13.5	13.5	12.6	Nov-10-04	4	314	4	t	Nov/11/04
12.7	12.3	12.5	12.7	12.6	13.1	13.5	13.5	12.6	12.9	Nov-11-04	5	315	4	t+1	Nov/12/04
12.3	12.5	12.7	12.6	13.1	13.5	13.5	12.6	12.9	13.9	Nov-12-04	6	316	4	t+2	Nov/13/04
12.5	12.7	12.6	13.1	13.5	13.5	12.6	12.9	13.9	13.9	Nov-13-04	7	317	4	t+3	Nov/14/04
12.7	12.6	13.1	13.5	13.5	12.6	12.9	13.9	13.9	13.8	Nov-14-04	1	318	4	t+4	Nov/15/04
12.6	13.1	13.5	13.5	12.6	12.9	13.9	13.9	13.8	13.3	Nov-15-04	2	319	4	t+5	Nov/16/04
13.1	13.5	13.5	12.6	12.9	13.9	13.9	13.8	13.3	12.8	Nov-16-04	3	320	4	t_6	Nov/17/04
13.5	13.5	12.6	12.9	13.9	13.9	13.8	13.3	12.8	12.2	Nov-17-04	4	321	4	t+7	Nov/18/04
13.5	12.6	12.9	13.9	13.9	13.8	13.3	12.8	12.2	13.6	Nov-18-04	5	322	4	t+8	Nov/19/04
12.6	12.9	13.9	13.9	13.8	13.3	12.8	12.2	13.6	13	Nov-19-04	6	323	4	t+9	Nov/20/04
12.9	13.9	13.9	13.8	13.3	12.8	12.2	13.6	13	13.1	Nov-20-04	7	324	4	t+10	Nov/21/04
13.9	13.9	13.8	13.3	12.8	12.2	13.6	13	13.1	13.1	Nov-21-04	1	325	4	t+11	Nov/22/04
13.9	13.8	13.3	12.8	12.2	13.6	13	13.1	13.1	13.3	Nov-22-04	2	326	4	t+12	Nov/23/04
13.8	13.3	12.8	12.2	13.6	13	13.1	13.1	13.3	13.3	Nov-23-04	3	327	4	t+13	Nov/24/04
13.3	12.8	12.2	13.6	13	13.1	13.1	13.3	13.3	13.3	Nov-24-04	4	328	4	t+14	Nov/25/04
12.8	12.2	13.6	13	13.1	13.1	13.3	13.3	13.3	13.2	Nov-25-04	5	329	4	t+15	Nov/26/04
12.2	13.6	13	13.1	13.1	13.3	13.3	13.3	13.2	13.1	Nov-26-04	6	330	4	t+16	Nov/27/04
13.6	13	13.1	13.1	13.3	13.3	13.3	13.2	13.1	12.3	Nov-27-04	7	331	4	t+17	Nov/28/04
13	13.1	13.1	13.3	13.3	13.3	13.2	13.1	12.3	13.3	Nov-28-04	1	332	4	t+18	Nov/29/04
13.1	13.1	13.3	13.3	13.3	13.2	13.1	12.3	13.3	13.2	Nov-29-04	2	333	4	t+19	Nov/30/04
13.1	13.3	13.3	13.3	13.2	13.1	12.3	13.3	13.2	13.5	Nov-30-04	3	334	4	t+20	Dec/01/04

Fig. 6. ANN structure regarding energy generation data for November, 2004

occurs in input layer, and how the data represent the episodic memory function at the network input.

There are nowadays several algorithms to train MLP networks. In the proposed work, the supervised back propagation algorithm is chosen. The parameter definition in training is not well understood sometimes, and it is called "black magic" [2]. Here, a learning rate of 0.75 and an inertial rate of 0.25 are used.

Training happens in the following way:

1) Initially, training data are selected from system data base. In this step, the selected data are called training set;
2) The training set samples are presented by the system to the ANN input layer, aiming to fulfill all the temporal window, that is, the FIR technique is used to incorporate the short-term memory effect in the ANN;
3) If the topology employs episodic memory (ANN 2), the system extracts data related to prediction period date: the day of the week, the day of the year, and the year;
4) With back propagation algorithm, the data propagation process occurs from input layer to the output layer. The output represents the prediction;
5) The outcome obtained from network output layer is compared to a real data of generated energy;
6) The error is back propagated, the weights between the units are adjusted, and the network operator is informed about the error rate found; and
7) If there is no training interruption, the input layer data are incremented in $t + 1$ and the cycle is repeated.

V. Tests and Results

After concluding training, the network is able to start the prediction process. The maximum period of prediction is undetermined; therefore one should consider that the ANN is trained in order to predict a short time period following the selection of training information; the larger the prediction sequence, the bigger the final error, because the prediction error is accumulated during the period [13].

In prediction process, the data are presented in temporal sequence to the ANN and the output neuron values are collected at the same training set period. Training happened with training data in the period $1^{st.}$ January, 2002 to $31^{st.}$ October, 2004, and the network prediction was obtained in the period $1^{st.}$ to $30^{th.}$ November, 2004, according to results shown in figure 7.

For the ANN 1, the obtained results have average error of 32.59%, comparing with the real result, since the ANN 2 has average error of 10.04%. Figure 8 shows the prediction results compared to generated energy in the period.

Prediction tests in other periods also show that the obtained results with ANN 2 are more realistic, that is, the network displays better performance in temporal series prediction. The episodic memories as well as short-term memory filter implementations have biological origin and confirm that the several existing divisions and their combinations can be successful in an ANN implementation.

Prediction is obtained by the following method:

1) The prediction time period is selected;
2) The last training sample is located in data base, its value is added as well as the number of samples referring to FIR filter in network input layer;
3) The system adds the obtained prediction sample to the one-dimensional array of input layer and shift the other data until losing the oldest datum;
4) If the topology employs episodic memory (ANN 2), the system extracts data related to prediction period date: the day of the week, the day of the year, and the year; and
5) The system repeats the cycle until the period selected by the operator is reached.

Figures 9 and 10 show two diagrams related to error prediction. Figure 9 displays the error rate obtained by the ANNs

Date	Real	ANN 1	Error (%)	ANN 2	Error (%)
1/11/2004	12,81	8,5335828	33,38%	12,90044203	0,71%
2/11/2004	12,65	10,641891	15,87%	13,08228869	3,42%
3/11/2004	12,3	10,820313	12,03%	12,83512183	4,35%
4/11/2004	12,52	9,457399	24,46%	13,17394519	5,22%
5/11/2004	12,66	11,150519	11,92%	13,1170589	3,61%
6/11/2004	12,56	11,597812	7,66%	13,03429299	3,78%
7/11/2004	13,1	10,997048	16,05%	13,15859314	0,45%
8/11/2004	13,48	10,368066	23,09%	13,12350031	2,64%
9/11/2004	13,47	9,696841	28,01%	13,17195744	2,21%
10/11/2004	12,63	8,8479533	29,94%	13,02128814	3,10%
11/11/2004	12,88	8,7710875	31,90%	13,12733754	1,92%
12/11/2004	13,94	9,4956595	31,88%	13,1099009	5,95%
13/11/2004	13,86	9,2172172	33,50%	13,11579627	5,37%
14/11/2004	13,77	11,0105	20,04%	13,05140854	5,22%
15/11/2004	13,31	10,253255	22,97%	12,98398703	2,45%
16/11/2004	12,75	6,5122173	48,92%	11,78536983	7,57%
17/11/2004	12,21	6,3992741	47,59%	10,79087151	11,62%
18/11/2004	13,63	6,3772668	53,21%	10,24055678	24,87%
19/11/2004	13,01	6,2375159	52,06%	10,48717756	19,39%
20/11/2004	13,14	6,3801269	51,45%	10,87627419	17,23%
21/11/2004	13,1	6,2720042	52,12%	10,81780526	17,42%
22/11/2004	13,28	6,3664599	52,06%	10,90049422	17,92%
23/11/2004	13,32	6,2941247	52,75%	10,79959075	18,92%
24/11/2004	13,26	6,2807698	52,63%	10,73084616	19,07%
25/11/2004	13,17	9,6420095	26,79%	10,64643328	19,16%
26/11/2004	13,11	9,1950776	29,86%	11,03232993	15,85%
27/11/2004	12,32	10,512239	14,67%	11,18135343	9,24%
28/11/2004	13,29	8,6217522	35,13%	11,23710423	15,45%
29/11/2004	13,16	8,565593	34,91%	10,90525395	17,13%
30/11/2004	13,51	9,326146	30,97%	10,80203962	20,04%
		AVG Error	32,59%	AVG Error	10,04%

Fig. 7. Prediction results of electric energy generation for ANN 1 and ANN 2, for November, 2004 and the average errors

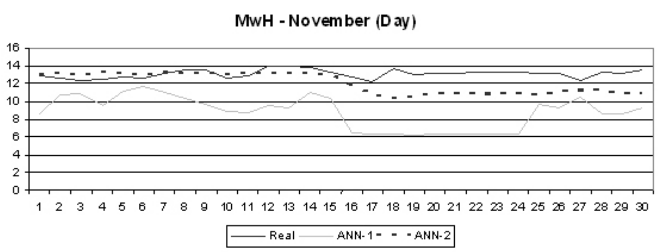

Fig. 8. Energy generation prediction result for November, 2004

in relation to generation prediction data of real energy in November, 2004, where the *y*-axis shows the error rate and the *x*-axis shows the days. Figure 10 shows the error rate evolution according to the sequential predictions. In this diagram, the *y*-axis shows the evolution through the accumulate error obtained from figure 9 and the *x*-axis shows the prediction days.

With these data, it is possible to observe the accumulative error through temporal series and it is possible to verify that the larger the prediction, the bigger the error will be at the end of the series. Although the prediction recommendation is only for 10 temporal series in trained resolution, the experiment was conducted with 30 series and, comparatively, it revealed a high average error:

- For ANN 1, that does not implement the episodic memory, it was obtained an average error of 32.59% in comparison with the real result for a 30-day prediction. If it is considered only the first 10 series, the average error is 20.24%.
- For ANN 2, that implements the episodic memory, the average error obtained is 10.04% for a 30-day prediction. For 10 days, this value is reduced to only 2.95%.

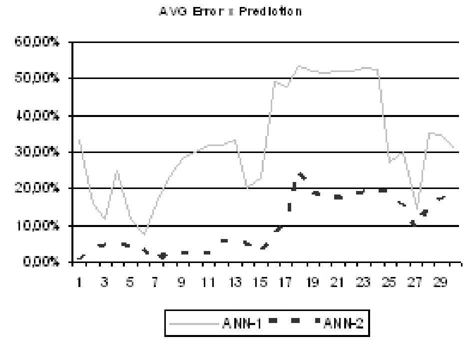

Fig. 9. Average error rate related to prediction

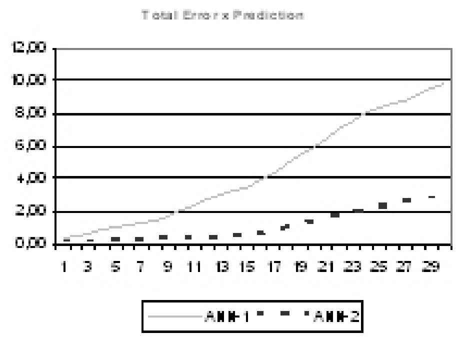

Fig. 10. Error rate evolution related to prediction

VI. Conclusions

This paper presents the combination of episodic long-term memories and short-term memories for temporal series prediction, by means of implementation of two topologies of artificial neural networks in order to predict electric energy generation in a small hydroelectric power station.

The implemented networks are of FIR type, and the use of episodic memory in one of the topologies displayed successful results, with a better efficiency compared to the network without episodic memory. The total average error for the applied case is 32.59% for conventional FIR network and only 10.04% for the FIR network with episodic memory.

The obtained outcomes with episodic memory implementation, for temporal series prediction, show that this type of memory is useful in applications that recover past experiences in order to anticipate future by means of predictions. They also show an increase in efficiency with the combined use with other types of memories, like the semantic memory, representing the network knowledge, and the short-term memory through FIR memory technique implementation.

Future research can be developed in this area, continuing the investigation of applications of other types of memories in artificial neural networks. It is also possible to utilize recurrent networks with episodic memory, especially in applications in which temporal series are essentially non-stationary.

In electric energy area, the proposal of a model enrichment consists in the inclusion, according to availability, of other hypothetical context variables like rain volume, etc., allowing a greater knowledge degree about the impact of these variables on generation potential of each station, resulting in more accurate future scenarios.

The model expansion in terms of applications in proper generation information scope, like active and reactive power demand prediction, and also the temporal order - hourly bases or load levels, considering larger horizons, can also be developed, once the sectoral exigency for such information already constitutes a reality nowadays.

REFERENCES

[1] Agência Nacional de Energia Elétrica - ANEEL, Documentos sobre o mercado de distribuição, 2004. Brazilian government: www.aneel.gov.br.

[2] M. A. Arbib, "Backpropagation: General Principles," in Michael A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks*, 2nd. ed. The MIT Press, Cambridge, Massachusetts, 2003.

[3] A. P. Braga and L. A. E. Soares, "Redes Neurais Artificiais e sua Aplicação em Sistemas Elétricos," *Revista Eletricidade Moderna*, número 233, Aranda Editora, São Paulo, SP, Brazil, 2001.

[4] M. L. Brandão, *As Bases Biológicas do Comportamento: Introdução à Neurociência*, Editora Pedagógica Universitária, 1ª Edição, Brazil, 2004.

[5] L. S. Caio and C. Bermann, "Análise das Metodologias de Previsão de Mercado de Energia Elétrica Face ao Novo Perfil de Planejamento no Ambiente Pós-privatização," in *Anais do III Congresso Brasileiro de Planejamento Energético*, São Paulo, SP, Brazil, 1999.

[6] S. H. Cardoso, "Human Memory: What it is and how to improve it," *Brain & Mind Electronic Magazine*, http://www.cerebromente.org.br/, n.1. 1997.

[7] J. L. Elman, "Finding Structure in Time," *Cognitive Science*, 14, pp. 179-211, University of California, San Diego, 1990.

[8] P. Gupta, "Word learning and verbal short-term memory: a computational account," in *Proceedings of the 18th. Annual Conference of Cognitive Science Society*, pp. 189-194, Hillsdale, NJ, 1996.

[9] S. Haykin, *Neural Networks - A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, New Jersey, 2 edition, 1999.

[10] I. Izquierdo, D. M. Barros, T. Mello e Souza, T., M. M. de Souza, L. A. Izquierdo, and J. H. Medina, "Mechanisms for memory types differ", *Nature*, n. 383, pp. 635-636. London, UK, 1998.

[11] R. Kandel, J. H. Schwartz, and T. M. Jessell, *Essentials of Neural Science and Behavior*, Appleton and Lange, Stamford, Connecticut, 1995.

[12] J. M. Rybash and B. E. Monaghan, "Episodic and semantic contributions to older adults' autobiographical recall," *The Journal of General Psychology*, January 01, 1999.

[13] D. Wang, "Temporal pattern processing," in Michael A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks*, 2nd. edition. The MIT Press. Cambridge, Massachusetts, 2003.