DYNAMIC SYSTEMS BASED ON NEURAL NETWORKS USED IN TIME SERIES PREDICTION

Professor Ph.D. **Valeriu LUPU**
Ştefan cel Mare University of Suceava, Romania
valeriul@seap.usv.ro

Associate Professor Ph.D. **Nina HOLBAN**
Ştefan cel Mare University of Suceava, Romania
ninah@seap.usv.ro

Abstract:

In this article the authors propose a modality of prognosis of the quantities of waste generated in a certain period. The proposal was finalized by achieving a model of prognosis by using dynamic systems based neural networks for the time series prediction. Time series with three components were used: trend, seasonality and residual variable. According to the input data, one can choose the adjustment model regarding the description of the phenomenon analyzed (additive and multiplying). In this scope the ***Cascade_Correlation*** *algorithm was used, a constructive learning algorithm. Starting from the input data a time series generates, with 1, 2 or 3 ahead (according to how we want to make the prognosis: for a month, for two months or for three months ahead). The advantages of the algorithm are the more rapid convergence and the elimination of the necessity to determine a priori the topology of the network. In the study the* ***Quickpropagation*** *learning algorithm was presented, used in order to involve the output units and candidate from the* ***Cascade_Correlation*** *algorithm. In the article a case study is presented for the analysis of data and for the time series prediction by using the soft made in Matlab. A comparison between the input data and those prognosticated by the neural network was made.*

Key words: algorithm, model of prognosis, quick propagation, neural network.

JEL classification: A10, C13

INTRODUCTION TO THE PRESENTATION OF THE DOMAIN

Modeling a neural network means taking into consideration some aspects of major importance in the functioning of the dynamic systems, of some parameters whose choice depends on the success or the failure of the network [2]. At present, there is no solid theory that is on the basis of a process of structural or functional implementation of a neural network in order to solve a certain problem, in reality this process is reduced to a trying-error type procedure, the process being rather close to art than to science.

Within the neural network implementation process, some aspects that must be taken into consideration are presented as follows:

- Pre- processing data
  - Data frequency (daily, weekly, monthly, etc.);
  - Data type;
  - Scaling method.
- Learning process
  - Start rate;
  - Moment term;
  - Tolerated error at learning;
  - Maximum number of rolls of the network;
  - Number of random initializations of the weights associated to connections.
  - Dimensions of the training, validation, testing sets.
- Topology of neural network
  - Number of neurons belonging to the input/ output layers;
  - Number of hidden layers;
  - Number of neurons from each hidden layer;
  - Type of activation functions;
  - Type of error function.

Taking these aspects into account, a method of implementation of neural networks in order to achieve the waste prognosis [1], [8] is presented in figure 1, with the mention that such a procedure may need to go repeatedly through previous stages, especially between the training and selection stage of the variables, following the previous-mentioned line of the trial- error process.

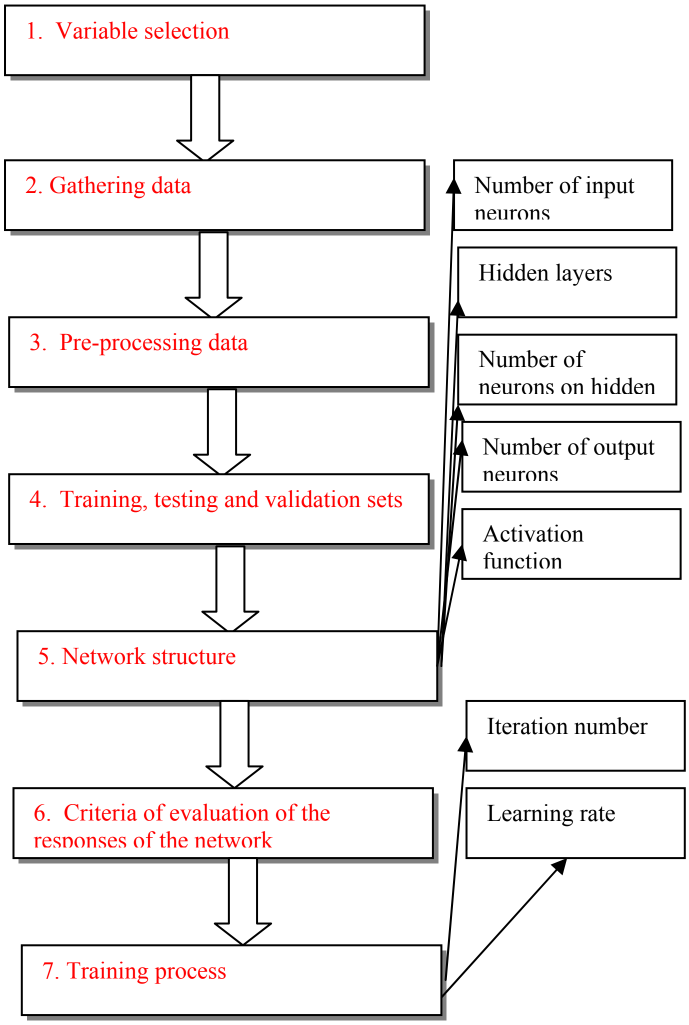


Figure 1. Implementation process of neural network

1.1. VARIABLE SELECTION

The success in the implementation of the neural network depends on the understanding and appropriate choice of the input variable. In the case of using the neural networks within the waste prognosis process, a support in this sense is obtained because of the theory supplied by the ecology domain [12],[14], which supplies a series of indicators that may be applied on the input data.

In case of achieving a prognosis regarding the time series, the network will have as a rule one output (so one neuron belonging to the last layer) supplying the prognosticated value, and the inputs may be represented through values of the variables analyzed at different previous moments. According to the type of prognosis, two modalities of specification of the output variable, and training of the network, respectively, can be distinguished:

- The network trained with one step ahead (Figure 3 a and b). It's about a prognosis of iterative type: the network will prognosticate the next value of the variable, value that it will use at the next step, etc. This process of prognosis may last endlessly and may be described like this:

$$X_{k+1} = \text{Neuralnetwork}\ (X_k, X_{k-1}, \ldots\ldots\ldots\ldots, X_{k-n})$$

❖ The network trained with 3 steps ahead (Figure 4 a and b). It's about direct prognosis which permits only to find one prognosticated value.

X_{k+3}=Neuralnetwork (X_k, X_{k-1},.............,X_{k-n})

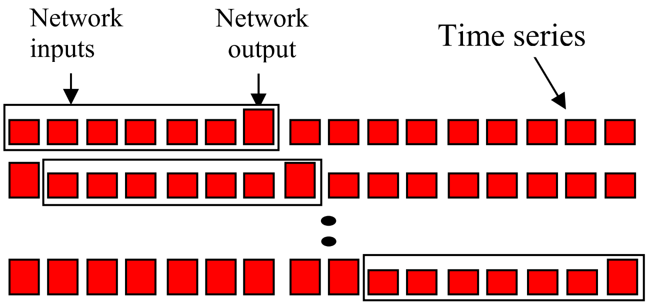


Figure 2. The values of the input and output variables in training with a step ahead

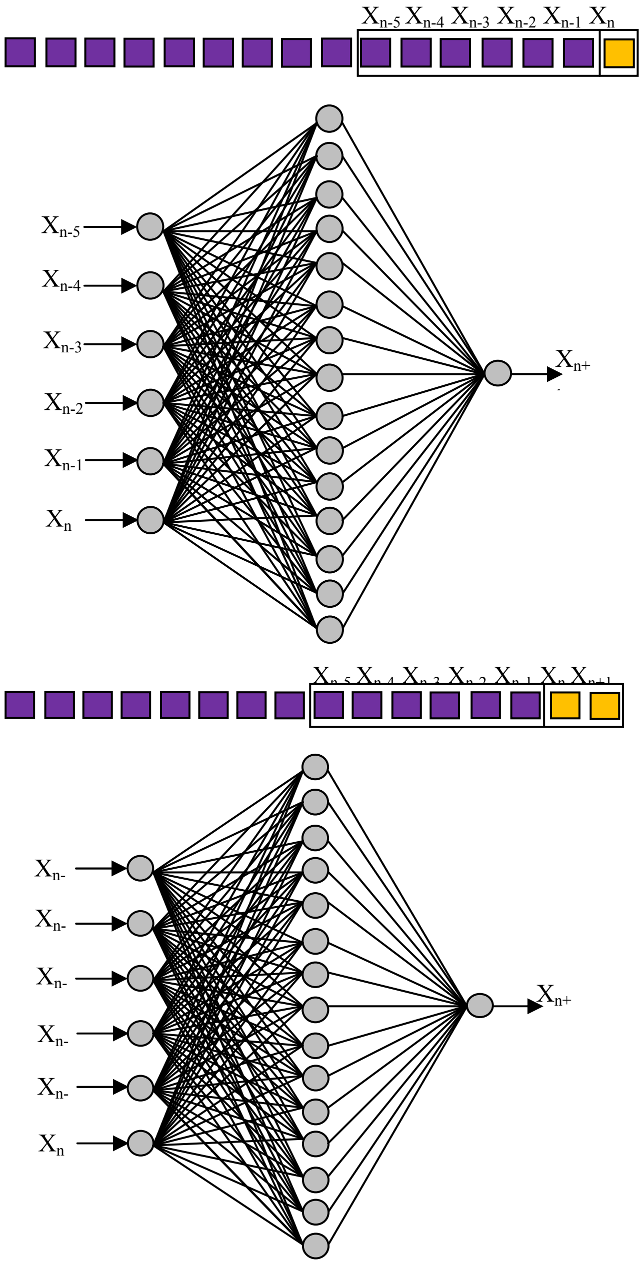


Figure 3.a. Prognosis with a step ahead for the first input vector; b. Prognosis with a step ahead for the last input vector

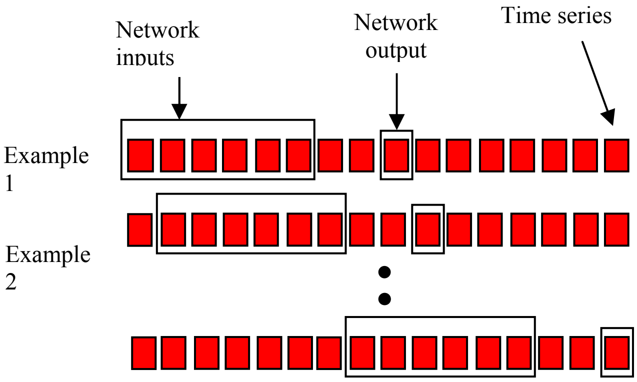


Figure 4.a. The values of the input and output values in training with 3 steps ahead

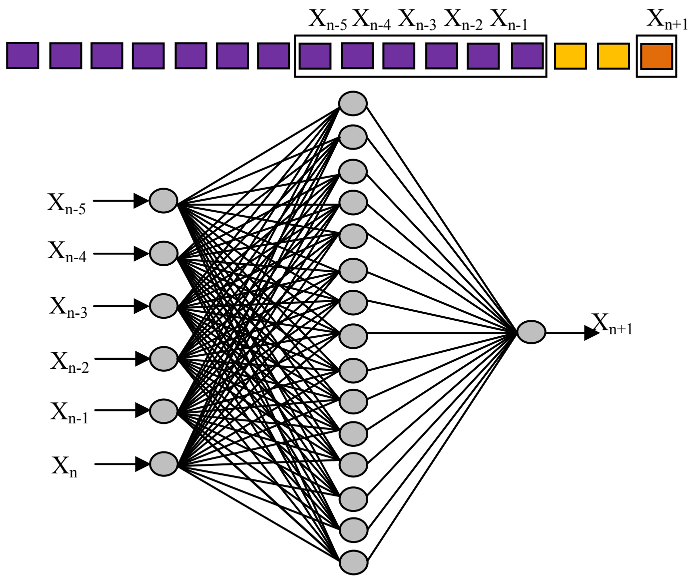


Figure 4.b. Prognosis with 3 steps ahead

2. NEURAL NETWORKS IN THE PROGNOSIS OF COLLECTED WASTE QUANTITIES

2.1. PROGNOSIS OF THE MANAGEMENT OF COLLECTED WASTE WITHIN A GIVEN PERIOD

In order to achieve the prognosis of the waste management a neural network of multilayer perceptor type (3 layers) will be used, whose architecture is presented in figure 5. The network is presented in figure 6. The network presents 6 neurons in the input layer, 15 neurons in the hidden layer and one input neuron, associated to the prognosticated value. A step ahead prognosis method is used, the output value may be written under the following functional form[2], [9],[12]:

$$y = x_{n+1} = f(x_{n-5}, x_{n-4}, x_{n-3}, x_{n-2}, x_{n-1}, x_n)$$

where $x_{n-5}, x_{n-4}, x_{n-3}, x_{n-2}, x_{n-1}, x_n$ represents the inputs of the network.

For the input data we will generate a time series as in table 1.

Year	1	2	3	4	5	6	7	8	9	10	11	12	Total	yi
2003	1420,1	1542,1	1522,8	1605,8	1704,1	1622,2	1622,3	1650,5	1607	1582,1	1566,6	1578	19023,6	1585,3
2004	1687,5	1745,1	1717,9	1734	1803,2	1752,6	1615,4	1700	1466,4	2200	1544,3	2800	21766,4	1813,87
2005	1571,9	1684,4	1684,2	2980	1631,9	1697,6	2600	1944,9	1915,7	3800	2057,8	2266,7	25835,1	2152,93
2006	2177,5	2091,6	2890	2246	2127,7	3200	4000	2900	3000	2071,6	2048,3	2106,6	30859,3	2571,61
2007	2021,3	2102,2	1922,6	2500	2092,2	5600	2900	2238,3	3000	5400	2242,8	5300	37319,4	3109,95
2008	2470,1	2473,6	4300	2505,3	4500	3400	2473,4	2543,8	2000	5600	2845,5	2888,7	38000,4	3166,7
2009	3400	3259,6	3429,1	4600	3562,7	3766,9	4405,5	3919,8	4154,9	4500	4900	4224,3	48122,8	4010,23
2010	4442,5	4693,9	5097,3	5107,44	5569,08	5897,44	5873,92	4833,89	4474,51	4671,12	5022,7	5002,39	60686,19	5057,18
2011	5159,96	4911,81	4884,2	5393,11	5069,83	5378,52	5101,87	5270,77	5149,83	5525,4	5896,04	6958,14	64699,48	5391,62
2012	6835,6	7644,55	7599,39	7414,68	7109,67	7600	7190,37	7244,79	6798,12				65437,17	7270,80
Total	31186,46	32148,86	35047,49	36086,33	35170,38	39915,26	37782,76	34246,75	33556,46	35350,22	28124,04	33124,83	411749,84	
yj	3118,65	3214,89	3504,75	3608,63	3517,04	3991,53	3778,28	3424,68	3356,65	3927,8	3124,89	3680,54		3520,69

Figure 5. The routine of testing the network by successively reading the data

The data are in a file data.txt with the structure $X_{n-5}, X_{n-4}, X_{n-3}, X_{n-2}, X_{n-1}, X_n$, X_{n+1} – where $X_{n-5}, X_{n-4}, X_{n-3}, X_{n-2}, X_{n-1}, X_n$ – represent the inputs of the network, and X_{n+1} represents the output of the network.

Table 1. Time series generated for input data

x_{n-5}	x_{n-4}	x_{n-3}	x_{n-2}	x_{n-1}	x_n	x_{n+1}
1420,1	1542,1	1522,8	1605,8	1704,1	1622,2	1622,3
1542,1	1522,8	1605,8	1704,1	1622,2	1622,3	1650,5
1522,8	1605,8	1704,1	1622,2	1622,3	1650,5	1607
1605,8	1704,1	1622,2	1622,3	1650,5	1607	1582,1
1704,1	1622,2	1622,3	1650,5	1607	1582,1	1566,6
1622,2	1622,3	1650,5	1607	1582,1	1566,6	1578
1622,3	1650,5	1607	1582,1	1566,6	1578	1687,5
....						
4884,2	5393,11	5069,83	5378,52	5101,87	5270,77	5149,83
5393,11	5069,83	5378,52	5101,87	5270,77	5149,83	5525,4
5069,83	5378,52	5101,87	5270,77	5149,83	5525,4	5896,04
5378,52	5101,87	5270,77	5149,83	5525,4	5896,04	6958,14
5101,87	5270,77	5149,83	5525,4	5896,04	6958,14	6835,6
5270,77	5149,83	5525,4	5896,04	6958,14	6835,6	7644,55
5149,83	5525,4	5896,04	6958,14	6835,6	7644,55	7599,39
5525,4	5896,04	6958,14	6835,6	7644,55	7599,39	7414,68
5896,04	6958,14	6835,6	7644,55	7599,39	7414,68	7109,67
6958,14	6835,6	7644,55	7599,39	7414,68	7109,67	7600

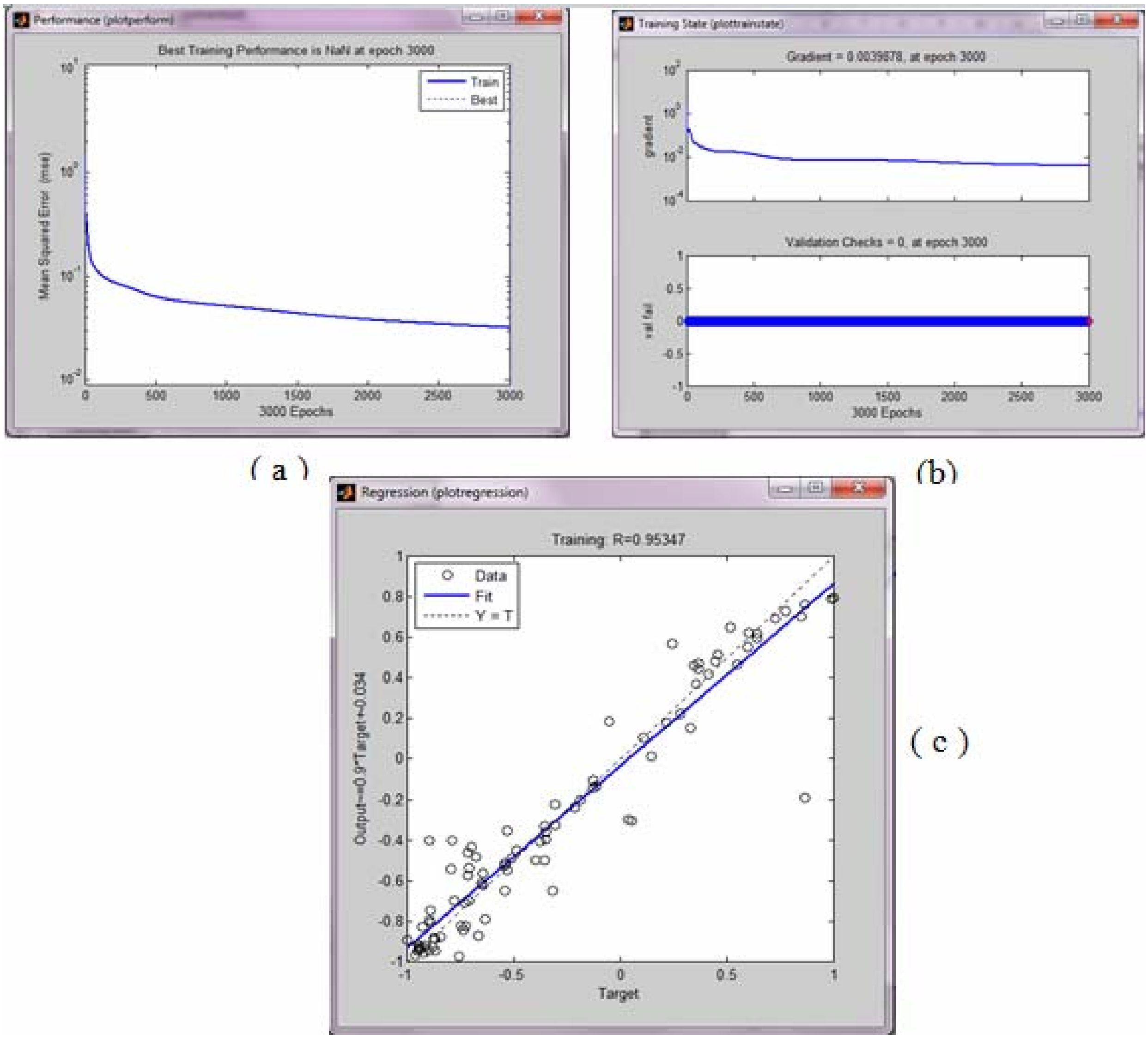
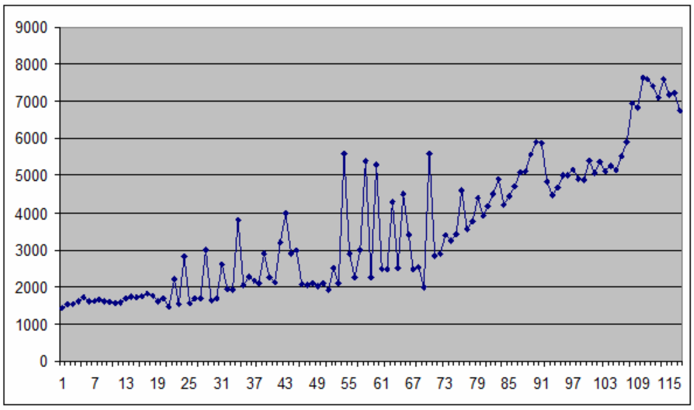
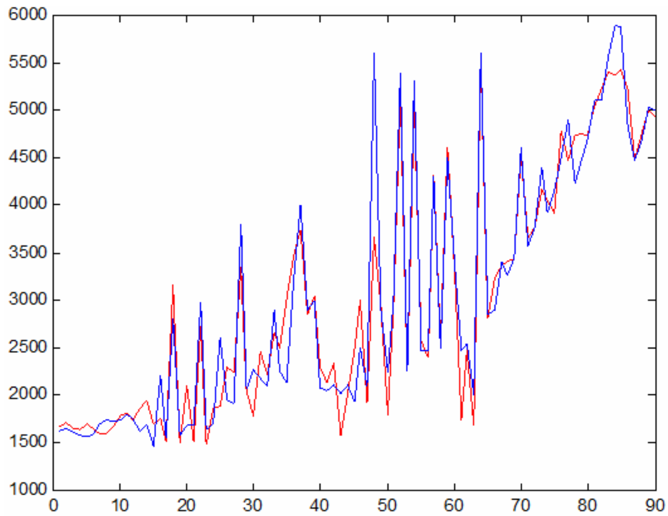


Figure 6. a) Performance of the network for 3000 periods b) Learning the network c) Analysis of regression of the network

The graphics of the input data and the graphic that compares the input data and those prognosticated are presented as follows (Figure 7a. and Figure 7.b).



(a)



(b)

Figure 7. a) The graphic for the input data (the values used for the training set (period 01/2003 – 06/2012)); b) Comparison between the input data (marked in red) and those prognosticated (marked in blue)

The data for the training of the network will be represented by the values from the period 01/2003 – 06/2012, so 114 values that make up the learning set [9], [12]. Among these 90 will be used for the effective training of the network, and the rest of 25 values will represent the test set. After achieving the training process, the network will be tested prognosticating the following 3 values of the next periods, 07/2012, 08/2012, 09/2012 (see Figure 8).

The pre-processing of the input data consists in their normalization, normalization achieved by the sigmoid function.

$$f(x) = \frac{1}{1 + e^{-\frac{x - \bar{x}}{\sigma}}}$$

Where: $\bar{x}$, σ represent the arithmetical mean, the deviation quadrangular mean respectively

$$\frac{\sqrt{\sum_{i=1}^{n_k}(x_i-\bar{x})^2}}{n_k}$$

(where nk represents the number of values making up the learning set (nk = 90)) of the values belonging to the training set. The values belonging to the training, testing set and the corresponding pre-processed values are to be found in table 2.

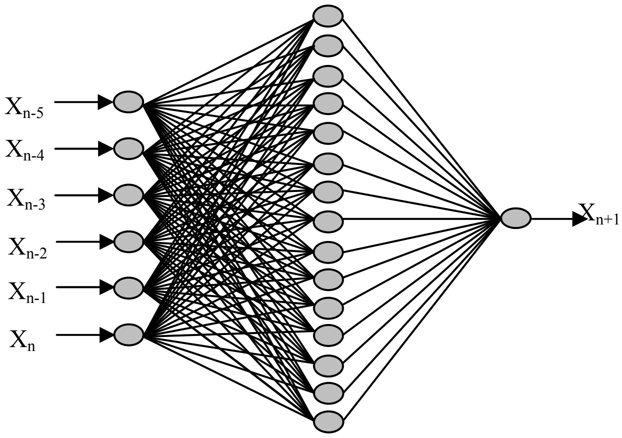


Figure 8. The network architecture for the prognosis of the quantities of generated waste. Neural network with 3 layers: 6 neurons belonging to the input layer, 15 hidden neurons and one neuron in the output layer

Table 2. Training set, testing set for the learning process

	1	2	3	4	5	6	7	8	9	10	11	12
2003	1420,1	1542,1	1522,8	1605,8	1704,1	1622,2	1622,3	1650,5	1607	1582,1	1566,6	1578
	0,499835	0,499845	0,499843	0,49985	0,499858	0,499851	0,499851	0,499853	0,49985	0,499848	0,499847	0,499848
2004	1687,5	1745,1	1717,9	1734	1803,2	1752,6	1615,4	1700	1466,4	2200	1544,3	2800
	0,499856	0,499861	0,499859	0,49986	0,499865	0,499861	0,499851	0,499857	0,499839	0,499896	0,499845	0,499944
2005	1571,9	1684,4	1684,2	2980	1631,9	1697,6	2600	1944,9	1915,7	3800	2057,8	2266,7
	0,499847	0,499856	0,499856	0,499958	0,499852	0,499857	0,499928	0,499876	0,499874	0,500022	0,499885	0,499902
2006	2177,5	2091,6	2890	2246	2127,7	3200	4000	2900	3000	2071,6	2048,3	2106,6
	0,499895	0,499888	0,499951	0,4999	0,499891	0,499975	0,500038	0,499951	0,499959	0,499886	0,499885	0,499889
2007	2021,3	2102,2	1922,6	2500	2092,2	5600	2900	2238,3	3000	5400	2242,8	5300
	0,499882	0,499889	0,499875	0,49992	0,499888	0,500163	0,499951	0,499899	0,499959	0,500148	0,4999	0,50014
2008	2470,1	2473,6	4300	2505,3	4500	3400	2473,4	2543,8	2000	5600	2845,5	2888,7
	0,499918	0,499918	0,500061	0,49992	0,500077	0,499991	0,499918	0,499923	0,499881	0,500163	0,499947	0,499951
2009	3400	3259,6	3429,1	4600	3562,7	3766,9	4405,5	3919,8	4154,9	4500	4900	4224,3
	0,499991	0,49998	0,499993	0,500085	0,500003	0,500019	0,50007	0,500031	0,50005	0,500077	0,500108	0,500055
2010	4442,5	4693,9	5097,3	5107,44	5569,08	5897,44	5873,92	4833,89	4474,51	4671,12	5022,7	5002,39
	0,500073	0,500092	0,500124	0,500125	0,500161	0,500187	0,500185	0,500103	0,500075	0,50009	0,500118	0,500116
2011	5159,96	4911,81	4884,2	5393,11	5069,83	5378,52	5101,87	5270,77	5149,83	5525,4	5896,04	6958,14
	0,500129	0,500109	0,500107	0,500147	0,500122	0,500146	0,500124	0,500138	0,500128	0,500158	0,500187	0,50027
2012	6835,6	7644,55	7599,39	7414,68	7109,67	7600	7190,37	7249,79	6738,12			
	0,50026	0,500324	0,50032	0,500306	0,500282	0,50032	0,500288	0,500292	0,500253			
Ma						3425,584						
σ						160,2947002						

For the same data Neuroshell [15] software was used for a three layer network: 6 neurons in the input layer, 15 neurons in the hidden layer and 1 neuron in the output layer. The time series formed from the input data (See Figure 6) shows like in Figure 9

Output neuron y = $\mathbf{x_{n+1}} = f(x_{n-5}, x_{n-4}, x_{n-3}, x_{n-2}, x_{n-1}, x_n)$

where $x_{n-5}, x_{n-4}, x_{n-3}, x_{n-2}, x_{n-1}, x_n$ represent the network inputs

(1420.1, 1542.1, 1522.8, 1605.8, 1704.1, 1622.2,	1622.3
(1622.3, 1650.5.1, 1607, 1582.1, 1566.6, 1578.0,	1687.5
(1745.1, 1717.9, 1734.0, 1803.2, 1752.6, 1615.4,	1700.0
(1466.4, 2200.0, 1544.3, 2800.0, 1571.9, 1684.4,	1684.2
(2980.0, 1631.9, 1697.6, 2600.0, 1944.9, 1915.7,	3800.0
(2267.7, 2177.5, 2091.6, 2890.0, 2246.0, 2127.7,	3200.0
(2900, 3000, 2071.6, 2048.3, 2106.6, 2021.3,	2102.2
(2500.0, 2092.2, 5600.0, 2900.0, 2238.3, 3000.0,	5400.0
(5300.0, 2470.1, 2473.6, 4300.0, 2505.3, 4500.0,	3400.0
(2473.4, 2543.8, 2000.0, 5600.0, 2845.5, 2887.7,	3400.0
(3259.6, 3429.1, 4600.0, 3562.7, 3766.9, 4405.5,	3919.8
(4154.9, 4500.0, 4900.0, 4224.3, 4442.5, 4693.9,	5097.3
(5107.44, 5569.08, 5897.44.0, 5873.92, 4833.89, 4474.51,	4671.12
(5022.7, 5002.39, 5159.96, 4911.81, 4884.2, 5393.11,	5069.83
(5378.52, 5101.87, 5270.77, 5149.83, 5525.4, 5896.04,	6958.14
(6835.6, 7644.55, 7599.39, 7414.68, 7109.67.4, 7600)→	7578.58

Figure 9. Time series with a step ahead implemented for the input data present in Figure 6

3. CONCLUSIONS

In the article neural networks for the prediction of time series corresponding to the quantities of waste collected in a certain period of time were used.

The advantages of using neural networks of Cascade – Correlation type for the prognosis of waste quantities for a given period (one months, two months or three months):

- Learning is supervised and very rapid;
- The network evolves;
- It starts from a simple (minimal) structure and the network structure evolves during learning by adding new neurons and new connections;
- It determines the structure (topology)
- The algorithm tries to eliminate 2 problems that appear at backpropagation and quickpropagation:
    - o The problem of the dimension of the step;
    - o The problem of moving target.

(a) The problem of the dimension of the step
   a. Small decreases of the gradient- local minimums;
   b. QP- because the approximation made the local minimum may be lost;

(b) The problem of the moving target:
   a. Each unit from the network is a characteristic detector;
   b. There is communication between the units of the network;
   c. Connects all the inputs of the network and the outputs of the hidden units;
   d. Maximizes the correlation between the intensification of input / hidden/ output units and the network error by updating the weights on the connections between these units;
   e. Chooses one or more candidate units with maximum correlation and freezes the weights on its inputs.

Finally it was demonstrated with NeuroShell software for solving the same problem. The solution is that the dynamic systems based on neural networks used for the prediction of time series, by using the quickpropagation algorithm, necessary to the training of output units and candidate from the Cascade – Correlation algorithm, uses constructive supervised learning, the convergence is more rapid and the elimination of the necessity to determine a priori the topology of the network.

BIBLIOGRAPHY

1. Steven Gonzalez, *Networks for Macroeconomic Forecasting: A Complementary Approach to Linear Regression Models*, 2000
2. Dumitrescu, D., Hariton, C., *Reţele neronale: Teorie şi Aplicaţii*, Ed. Teora, Bucureşti 1996
3. Geva, A., ScaleNet – *Multiscale neural Network Architecture for Time Series Prediction, IEEE Transactions on Neural Networks*, 9(5), 1471 – 1482
4. Masters, T., *Practical Neural Network Recipes in C++*, academic Press, Boston 1993
5. Rumelhart, D.E., Hinton, G.E., Williams, R.J., *Parallel Distributed Processing*, MIT Press, Cambridge, MA, 1986
6. Tacu, A.P., Holban, St.,Vancea, R., Burciu, A., *Inteligenta artificiala: Teorie si Aplicatii in Economie*, Ed. Economica, Bucuresti, 1998
7. Wasserman, P.D., *Neural Computing. Theory and Practice*, Van Nostrand Reinhold, New York, 1989
8. Leustean L, *Applications on neural networks to some echological phenomena forecasting, The Proceedings of the 3nd International symposium of Economic Informatics*, Bucharest, 1997
9. Brockwell Peter, *Introduction to time series and forecasting*, New York, 2000;
10. Basheer I.A., *Artificial neural networks, fundamentals, computing, design and application*, Journal of Microbiological methods, Elsevier Science, 2000
11. Yu Hen Hu, Jeng-Neng Hwang*, Handbook of Neural Networks. Signal Processing*, CRC Press, 2002
12. Lupu Valeriu, Lupu Catalin, Morariu Nicolae, *The implementation of clean production and the use of neural networks in forecating waste management*, WSEAS Transactions on Systems and Control, Issue 9, Vol.3, 2008
13. Marie-France Paquet and Timothy C. Sargent *Forecasting employment Rates: A Cohort Approach*, 2000
14. Brachinger, H.W., M. Weber, *Risk as a Primitive: a Survey of Measures of Perceived Risk,* ORSpektrum, Spinger-Verlag, Berlin, 2007.
15. Kriangsiri Malasri, et.all, Backpropagation Network for Time Series Forecasting, NeuroShell 2, Release 4.0. Ward Systems Group, Inc., Executive Park West, 5 Hillcrest Dr., Frederick, MD 21703, www.wardsystems.com, 2000.