Dipartimento di Scienze Economiche, Matematiche e Statistiche

Università degli Studi di Foggia

Forecasting Photovoltaic Deployment with Neural Networks

Crescenzio Gallo, Michelangelo De Bonis

Quaderno n. 02/2011

Dipartimento di Scienze Economiche, Matematiche e Statistiche, Largo Papa Giovanni Paolo II, 1,
71121 Foggia (Italy), Phone +39 0881-75.37.30, Fax +39 0881-77.56.16

Forecasting Photovoltaic Deployment with Neural Networks

Crescenzio Gallo, Michelangelo De Bonis
IEEE Members
[*c.gallo, m.debonis*]*@ieee.org*

Dipartimento di Scienze Economiche, Matematiche e Statistiche
Università di Foggia
Largo Papa Giovanni Paolo II n.1, 71121 Foggia, Italy
Phone +39 0881-753708 Fax +39 0881-753709

Abstract

The photovoltaic (PV) industry in Italy has already crossed the threshold of 1 GW of installed capacity. Currently there are approximately 70,000 certified facilities in operation for a power generation of 1,300 GWh/year. With these figures, Italy has become the second country in Europe for PV installed power after Germany.

The energy produced would be sufficient to meet the power needs of approximately 1,200,000 people. This leads to some questions: Will this technology continue to grow exponentially even after the recent reduction in rates by the Energy Bill? Will the number of installed PV facilities still grow even with less public support and (probably) a reduction in the technology purchase price?

The purpose of this paper is therefore to develop a conceptual model to make a prediction of the PV installed power in Italy through the use of "supervised" artificial neural networks. This model is also applied to the analysis of the spread of this technology in some other European countries.

Keywords: photovoltaic, forecasting, neural networks.
ACM Taxonomy: 3.I.III.IX

Contents

1 Introduction

In this work we want to develop and apply a computing model for forecasting the future deployment of one of the sustainable electricity options, solar photovoltaic (PV) technology in Italy. The conceptual model is developed under the assumption of PV modules widely manufactured in the market at present (see Figure 1), and the future implications of using PV technology in the electricity sector is evaluated.

The word *sustainable* in this context implies energy, environmental and economic sustainability. Generating cleaner electricity when compared to the grid electricity sources constitutes environmental sustainability. PV electricity mitigates CO2 emissions from the grid. Inclusion of such monetary ben-

Figure 1: An installation of PV panels

efits from CO2 mitigation into the evaluation of the economic performance of PV technology should encourage economic sustainability.

After a first assessment of the state of the art in Italy, we examine the motivations at the base of the present work. Then, a conceptual model is analyzed and developed through a supervised artificial neural network, followed by some experimental results. Finally, a series of considerations are made to focus on the current research and the future directions.

2 State of the art of PV in Italy

A study by the Politechnic of Milan analyzed the state of the art of photovoltaics in Italy in 2009, highlighting its characterization, the prospects for development and potential in the Italian market.

The total photovoltaic power installed in Italy in 2009 amounted to approximately 490 MW, and during 2010 it crossed the threshold of 1 GW of installed capacity. According to an approach that takes as reference the installed capacity per capita in Italy there is an average installed power of 10.3 kW every 1,000 inhabitants. Currently there are approximately 70,000 certified facilities in operation for a power generation of 1.3 GWh/year. With these figures, Italy has become the second country in Europe for PV installed power after Germany.

The total amount of photovoltaic capacity, both off-grid and on-grid, to be installed (in Italy and other countries) is expected to increase in the future through 2020. The PV development forecasts within 2020 in Spain and Germany respectively reach 651 and 865 kW per 1,000 inhabitants, almost one kWp per capita. In order to bridge the current photovoltaic gap, photovoltaic installations in Italy could be reasonably estimated at around 45 GW by 2020: about 0.75 kWp installed per capita.

2.1 The Energy Bill

On June 5, 2009 was published the Directive 2009/28/EC of the European Parliament and the Council on the promotion of energy from renewable sources. The measure marks a major turning point in setting the EU energy policy in that, for the first time, the theme of renewable energy is faced with a global vision.

With Directive 2009/28/EC, the Community has set itself the goal of meeting by the year 2020, a share of at least 20% of final energy consumption by using renewable sources. Member States were therefore assigned binding targets which, unlike those established under the previous regime, are not attributable to individual policy areas (e.g. production of electricity, use of biofuels for transport), but embrace across all types of use of energy products.

The strategies to be adopted at national level in order to attain the objective set for Italy – 17% coverage of the final consumption of energy through renewable sources by 2020 – must therefore take into due consideration the general character of the new Community measure. It will be necessary to act in a coordinated manner and to reduce consumption, to achieve a full exploitation of the use of renewable sources to satisfy power consumption, heat and transport sector.

In anticipation of a Directive of the European Parliament, the Ministry of Economic Development and the Ministry for the Environment, Land and Sea issues the Decree of 19.02.2007 "Criteria and methods for increasing the production of electricity by photovoltaic conversion of solar source..." confirmed to the Manager of Electrical Services – GSE SpA – the role of implementing the incentive mechanism of the photovoltaic known as the "Energy Bill".

The plants came into operation after 01.01.2010 are entitled to an incentive rate paid for a period of twenty years – from the date of entry into the facility – which remains in constant currency for the entire period.

The higher rates are approved for the small household systems of up to 3 kW which are architecturally integrated. The lowest rates are valid for large systems which are not architecturally integrated. Rates are provided for a period of twenty years from the date of entry into operation into the facility and remain constant, that is not subject to ISTAT updates, for the entire period. The values in the above table were calculated with a deduction of 4% rates reported in the Min. Decree of 19.2.2007 (2% for each year subsequent to 2008). Then we are witnessing the spread of this technological innovation in the social system.

3 Motivations

With such increasing fraction of PV electricity in the grid resource profile in the future, the primary motivation of this research arises from the need for examining certain implications of generating increased PV electricity in Italy in the future. The front end implications include primary energy, cost, labor consumption, and environmental impacts associated with the use of different types of PV technologies. PV panels generate different amounts of electricity based on the solar radiation available at various locations. Photovoltaic electricity does not displace the entire average mix of resources in the grid [1]. Hence there is a need to develop methodologies to accurately estimate the potential CO2 abatement deriving from installed PV electricity at peak demands.

Increased PV electricity generation has significant economic implications. The cost of PV electricity has decreased from \$5.4 per Wp (in 2001) to \$4.8 per Wp (2009) [2]. With increased installation in the future, one of the motivations is also to evaluate the specific technology and policy changes that will facilitate the highest increase in the economic performance of PV technology. In the future, the increased deployment of PV technology cannot be evaluated in isolation but in competition with fossil based, non renewable and other renewable technologies. The PV deployment under such a competitive scenario is indeed dependent on its decreased production cost (due to learning curve and economies of scale effect) and CO2 emission factor.

Hence evaluating the amount of PV electricity to be generated in the future under constraints of a CO2 cap is also an important motivation for this research.

In this context we studied the effect of the deployment of photovoltaic panels in the production of clean energy through the use of mathematical models measuring the amount of clean energy produced, which could be used at a forecasting level for strategic planning (e.g. for the modulation of the incentive fund) and/or investment control and feed-back.

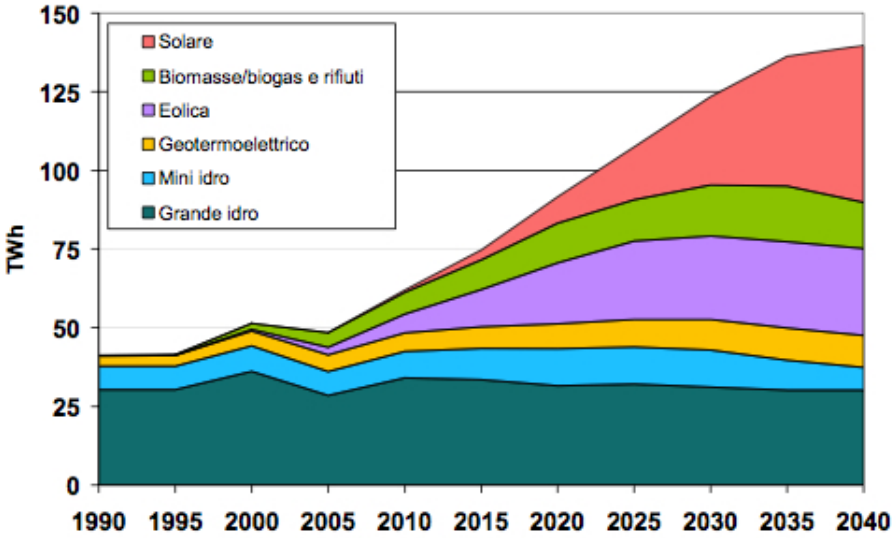


Figure 2: The contribution of different renewable sources in the acceleration technology scenario (Source: ENEA)

The adoption of technological innovations such as photovoltaics to produce clean energy on a large scale within a social system would solve the problem of minimizing emissions in energy production. This is a topic of great importance because, according to the prevailing valuations, it is important to reach certain levels in good time to tackle the huge growth in energy demand from Asian countries holding large reserves of coal (see Figures 2 and 3).

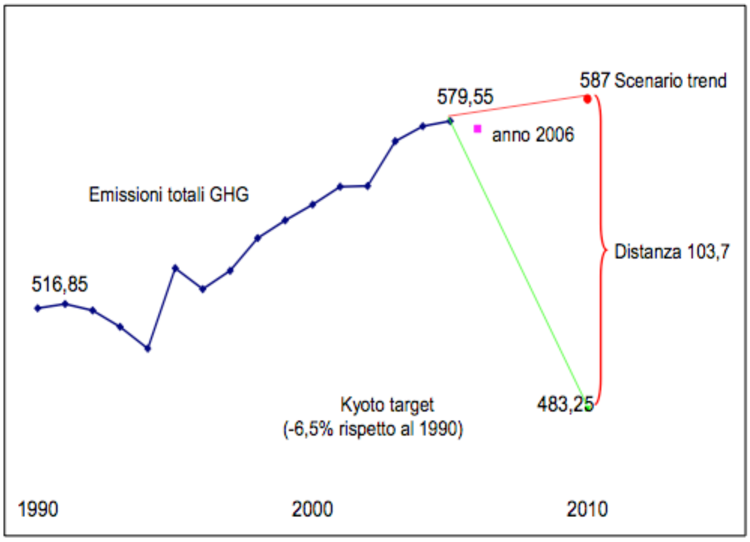


Figure 3: Emissions and the Kyoto target point assessment for 2010 (Mt CO2 eq.) (Source: ENEA)

4 The Mathematical Model

The mathematical model was developed after the following time discrete assumption:

$$y_t = y_{t-1} + g(t)(m - y_{t-1})$$

where

y_t	cumulative installed capacity to be forecasted at time t, in MW;
y_{t-1}	cumulative installed capacity at time $t-1$, in MW;
$g(t)$	diffusion coefficient;
m	maximum installable PV power;
$(m - y_{t-1})$	residual installable PV power.

The diffusion coefficient $g(t)$ is still the sum of two terms: the attraction function $h(t)$ for the purchase of photovoltaic, and the incentive mechanism introduced by governments for the installation of photovoltaic sites:

$$g(t) = \alpha y_{t-1} h(t) + \beta \sqrt{y_{t-1}}$$

where:

α	process growth rate constant, representing new PV installations as a fraction of cumulative installed capacity until t;
$h(t)$	attraction function for buying PV;
$\beta \cdot \sqrt{y_{t-1}}$	a factor related to the incentives introduced by government to stimulate PV new installations.

The attraction function $h(t)$ can be assumed as the difference of two costs:

$$h(t) = c^{NR} - c_t^{PV}$$

where c^{NR} is the cost of one kWh produced by a non-renewable source of energy and c_t^{PV} is the cost of one kWh produced by a photovoltaic system.

The discrete mathematical model developed for the prediction of growth of “on-grid connected” photovoltaic systems in Italy, at the base of this work, so is the following:

$$y_t = y_{t-1} + y_{t-1}\left[\alpha(c^{NR} - c_t^{PV}) + \frac{\beta}{\sqrt{y_{t-1}}}\right](m - y_{t-1})$$

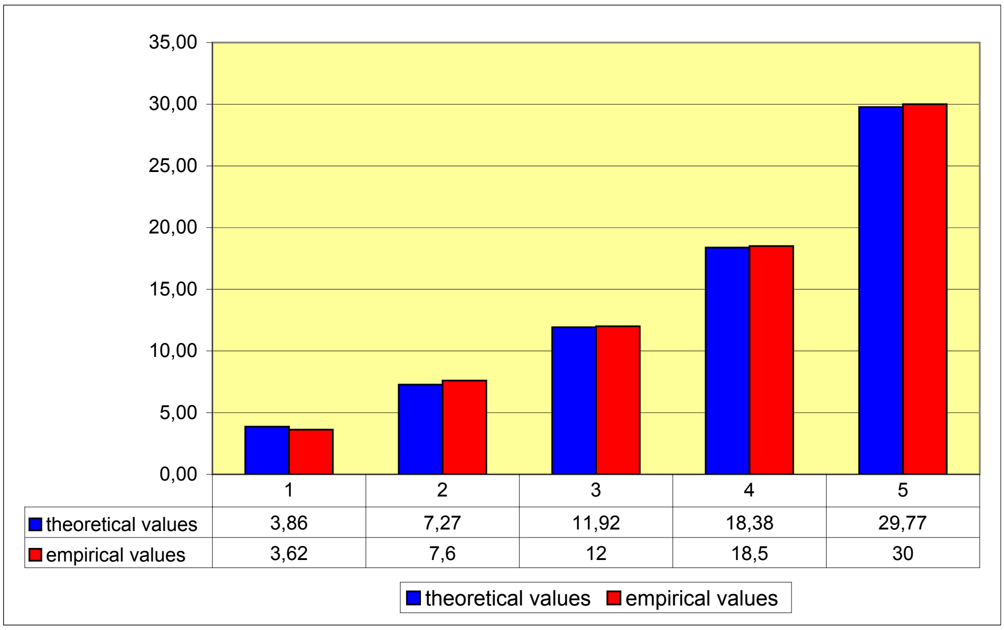
where it should be noted that the computation of the trend of cumulative PV power at time t is due to certain factors including, above all:

- the development of previously cumulated power;
- the incentive mechanism introduced by governments to promote the installation of photovoltaic systems in the area;
- the cost to produce one kWh from a PV system.

The results of its analytical application to the available data are summarized in Table 1, where installed powers are expressed in MW.

Table 1: Forecasting results in MW obtained from the mathematical model

YEAR	2002	2003	2004	2005	2006
Empirical	3.62	7.60	12.00	18.50	30.00
Theoretical	3.86	7.27	11.92	18.38	29.77



The results are summarized by the value of the square index of deviation $I_2 = 0.012204$, independent of the units used to express the data on which it is applied and computed as follows:

$$I_2 = \frac{\sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}}}{\frac{\sum_{i=1}^{n} x_i}{n}}$$

where $x_1, \ldots, x_n$ are the theoretical data, $y_1, \ldots, y_n$ the empirical data and n is the total number of observations.

The model has subsequently been tested through the use of a supervised neural network[1], confirming its validity and highlighting the significant contribution that can be obtained with such tools.

5 The Neural Network Model of the System

5.1 Dynamical discrete systems

The techniques used in this work are the classical for evaluating a system process by a series of observed data, and fall under the general category of

---

[1] Implemented in the Wolfram Mathematica software environment available in the Laboratory for the Quantitative Analysis of Data of the Dept. of Economics, Mathematics and Statistics of the University of Foggia.

system identification. Figure 4 illustrates the concept of discrete dynamical system underlying the assumed model.

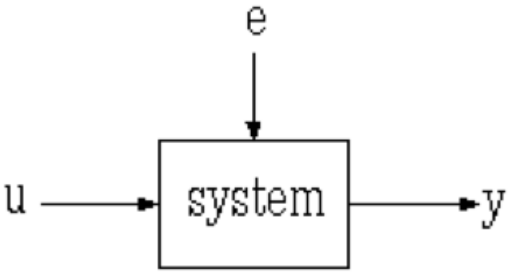


Figure 4: A dynamical system with input signal $u(t)$, noise signal $e(t)$ and output signal $y(t)$.

The output signal $y(t)$ from system is observable and measurable, and it is the signal we want to understand and describe. The input signal $u(t)$ is an external measurable signal influencing the system. Finally, the noise signal $e(t)$ affects the system but, unlike the previous signal, it is not measurable.

All these signals are time dependent.

5.2 Modeling PV growth through a neural network

To simulate the growth of PV we consider a discrete dynamic model in which the cumulated PV power at time t is the sum of two factors:

- the previous PV cumulated power;
- the diffusion coefficient multiplied by the still installable photovoltaic power.

The dynamical system considered above can also be modeled by an artificial neural network. This network consists of a combination of *FeedForward* or *Radial Basis Function* neural networks, and a specification of the vector of inputs to the network. Both of these parts must be specified by the user.

The input vector, or vector of regressors (as is often called talking about dynamical systems), contains the values of past input and output values of the system specified by three indices: n_a, n_b and n_k. So the shape of the input vector for the model of the dynamic system can be written as follows:

$$x(t) = [y(t-1) \cdots y(t-n_a) u(t-n_k) \cdots u(t-n_k-n_b+1)]^T$$

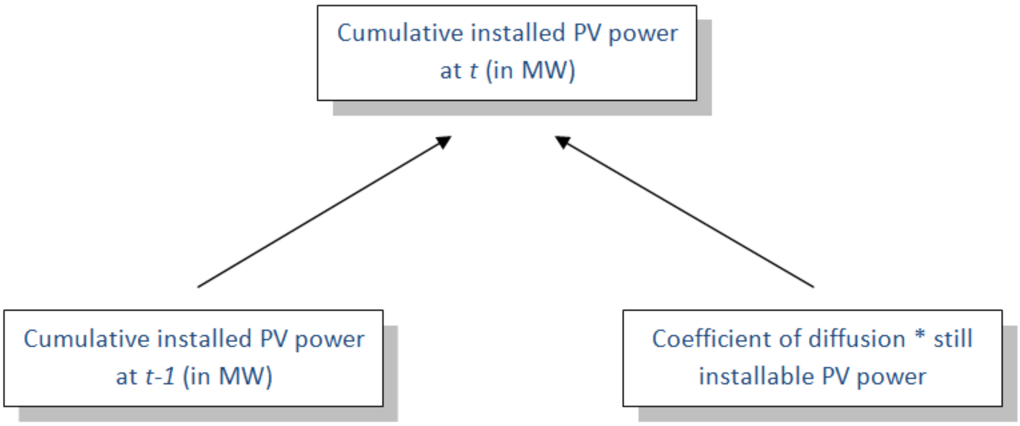


Figure 5: The cumulative installed PV power at time t.

Index n_a represents the number of past output values that are inserted as input of our time series, also known as "order of the model". Value n_b represents the number of past input values taken as inputs, and finally n_k represents a simple displacement of the temporal sequence of input values to be entered into the system.

A models with regressors, as in the expression previously reported, is called ARX model (Autoregressive model with eXtra input signal). Figure 6 shows a neural network ARX model with a layer of hidden neurons and a feedforward type network (in fact there are no cycles among the various network elements).

6 Implementation of the Neural Network

6.1 Input data

The data on which to test the neural network (shown in Figure 7 as input to Mathematica) are the same on which the discrete mathematical model has been tested. They come from the document entitled "National Survey Report on PV Power Applications in Italy 2009" provided by the IEA (International Energy Agency). The annual cumulative power, available from 1992 to 2009, is expressed in MW. Price refers to the production cost and is expressed in Euro per W.

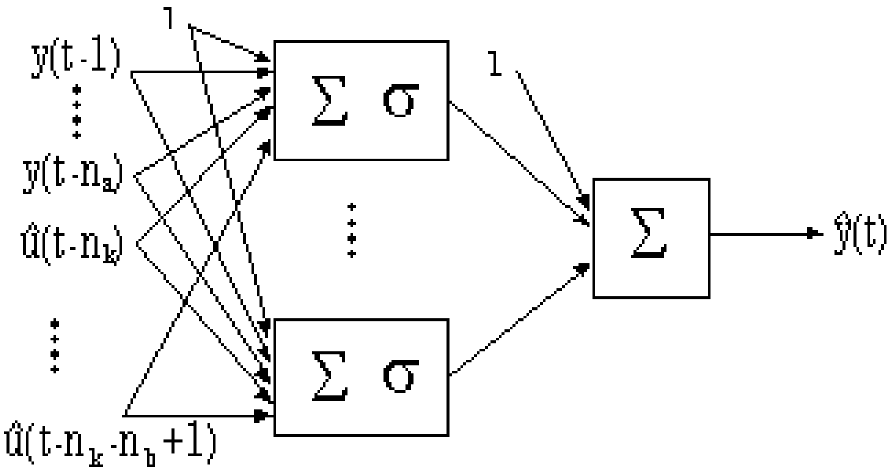


Figure 6: An ARX neural network model

```
<<Italia.dat
anno = {
{1992},{1993},{1994},{1995},{1996},{1997},{1998},{1999},{2000},
{2001},{2002},{2003},{2004},{2005},{2006},{2007},{2008},{2009}
}
potenza = {
{0.100},{0.100},{0.150},{0.335},{0.404},{0.677},{0.780},{0.905},
{1.155},{1.635},{3.620},{7.600},{12.000},{18.500},{30.500},
{83.900},{295.000},{656.800}
}
prezzo = {
{9.27319},{9.64063},{9.32129},{9.32129},{9.32129},{9.32056},
{9.3},{8.8},{7.75},{7.4},{7.5},{7.3},{6.8},{7},{6.4},{6.5},{6},{4.5}
}
```

Figure 7: Input data to the neural network in Mathematica list format

6.2 The neural network model

For creating the neural network model a series of input data were used, including the vector of prices (on which the neural network will have to practice), the yearly data series of the cumulated power, the vector of regressors n_a, n_b, n_k (which determines the number of inputs to the neural network and the time horizon used for the forecast) and the number of hidden neurons n_h.

The Mathematica code used to create the neural network and the related computational error are shown in Figure 8.

```
{model1,fitrecord} = NeuralARXFit[prezzo, potenza,
                        {{1},{2},{0}},FeedForwardNet, {2,2}];
```

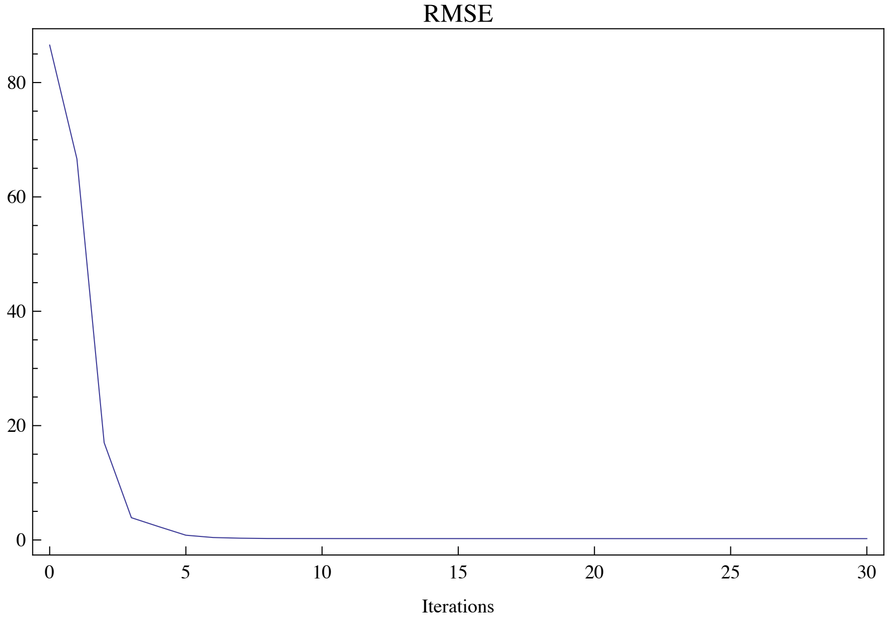


Figure 8: Performance of the round mean square error (RMSE) in creating the neural network model.

The creation of the neural network performed by the code described above results in the model shown in Figure 9 where in input we have:

- the power cumulated at the immediately previous time;
- both the current and the previous price.

The two neurons in the hidden layer are fully connected to the inputs and participate in the estimated output. Their activation (sigmoid) function is typical of the problems of “time series”.

`NetDisplay[model1]`

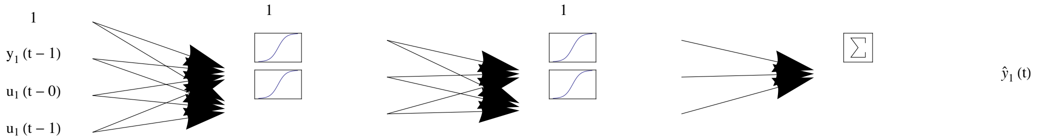


Figure 9: Visual representation of the neural network model.

The mathematical form corresponding to the mathematical model generated by the neural network is as follows:

$$\begin{aligned}\{&-141.704/(\exp(-9.98013+44.7513/(e^{0.00487548\,p(t-1)+0.048626\,p(t)+0.0143083\,y(t-1)-0.298016}+1)-\\&14.0877/(e^{-0.0021455\,p(t-1)-0.028486\,p(t)-0.00888827\,y(t-1)-0.748961}+1))+1)-\\&15073.7/(\exp(-8.71433+4.04657/(e^{0.00487548\,p(t-1)+0.048626\,p(t)+0.0143083\,y(t)-1-0.298016}+1)+\\&12.8828/(e^{-0.0021455\,p(t-1)-0.028486\,p(t)-0.00888827\,y(t-1)-0.748961}+1))+1)+1110.5\}\end{aligned}$$

where

- $p(t-1)$ corresponds to the price at time $t-1$;
- $p(t)$ is the price at time t;
- $y(t-1)$ represents the comulated power at time $t-1$.

6.3 Theoretical vs. empirical data

The chart in Figure 10 shows the trend curves of empirical data, i.e. those actually observed, and theoretical data, namely those forecasted by the neural network. Note that, except for an initial phase of adaptation of the neural network to the data, the graph shows a good predictive power of the model since the two curves, with increasing time, overlap.

```
NetComparePlot[prezzo, potenza, model1, PredictHorizon->1,
      PlotStyle-> {Hue[.0],Hue[.6]},
      PlotLegend-> {"Empirico","Teorico"}]
```

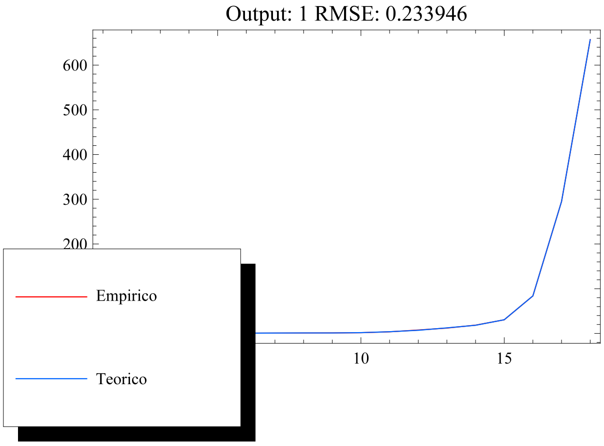


Figure 10: Representation of the trend of empirical vs. theoretical data.

6.4 Hidden neurons

Some parameters of the neural network can be represented and examined to better understand the behavior of the model. In Figure 11 the graph shows the trend, as a function of time, of the hidden neurons when the model is used to predict the data.

6.5 Linear parameters

The graph in Figure 12 shows the parameters, derived from the linear model at each time, of the regression vector vs. the analyzed data.

6.6 Final error distribution

Of utmost importance is also the analysis of the errors made by the model in making the prediction (in our case the cumulated powers) on the basis of available data.

The histogram in Figure 13 is the representation of the result of the error distribution of the model applied to the input data.

```
NetPlot[model1,prezzo,potenza,DataFormat-> HiddenNeurons,
   PlotStyle-> {Hue[.1],Hue[.3],Hue[.5],Hue[.7],Hue[.9]},
   PlotLegend-> Map["Neurone N\[Degree] "<>ToString[#]&,
   Range[Length[model1[1,1,1,1]]]]]
```

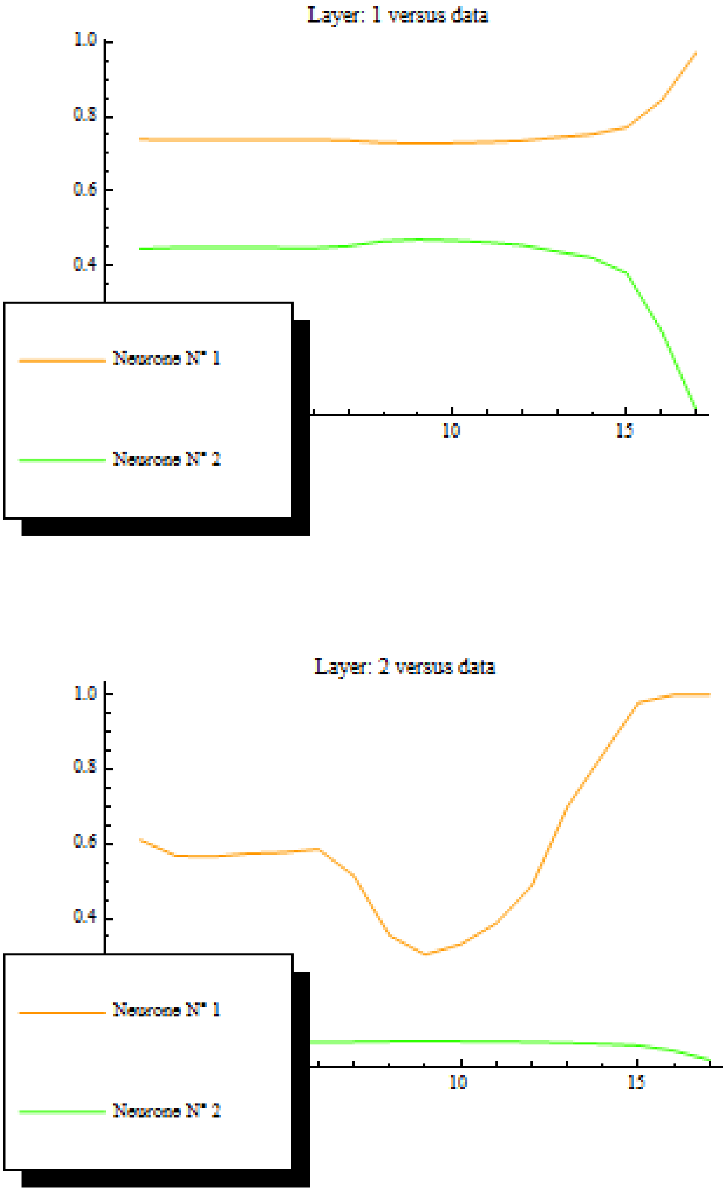


Figure 11: Hidden neurons' values in the neural network model.

```
NetPlot[model1,prezzo,potenza,DataFormat-> LinearParameters,
      PlotStyle-> {Hue[.1],Hue[.5],Hue[.9]}]
```

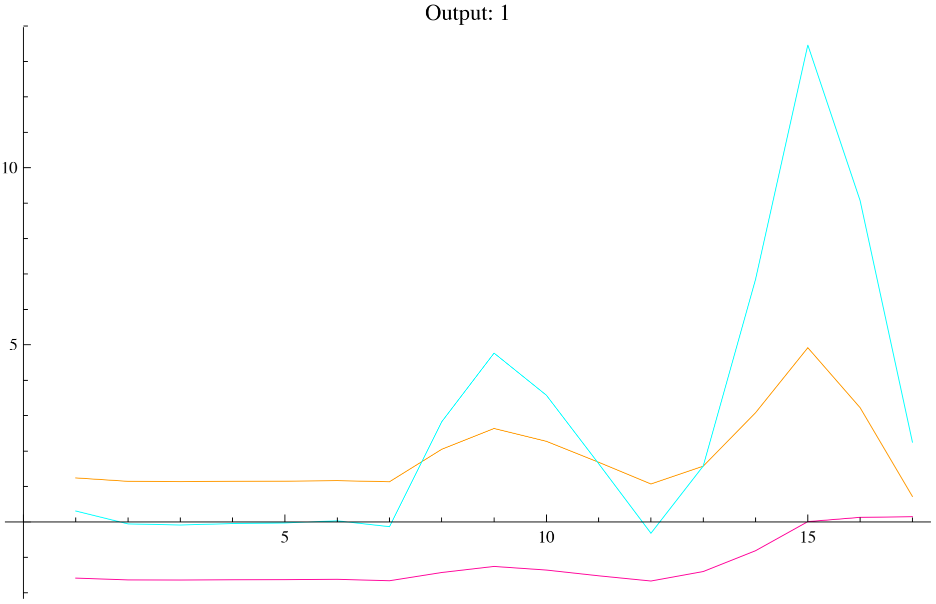


Figure 12: Representation of the linear parameters of the neural network.

```
NetPlot[model1,prezzo,potenza,DataFormat-> ErrorDistribution]
```

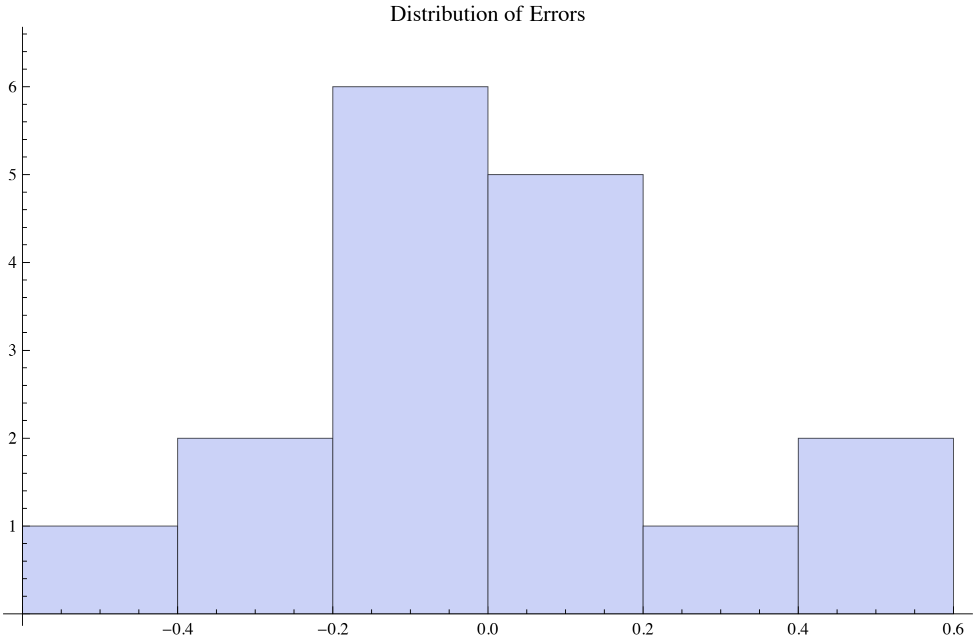


Figure 13: Error distribution of the neural network model.

6.7 Forecasting results

To obtain estimates of the real cumulated powers, we have applied to the neural network model developed earlier the data vector inputs, the cumulated power at the previous year, the price per Watt produced in the current forecasting year and in the previous one:

```
model1[[1]][{1.635, 7.5, 7.4}] (* y (2002)=3620 KW *)
{3.80}
model1[[1]][{3.798123572024224‘, 7.3, 7.5}] (* y(2003)=7600 KW *)
{7.34}
model1[[1]][{7.3408999441937794‘, 6.8, 7.3}] (* y(2004)=12000 KW *)
{12.14}
model1[[1]][{12.138557099791342‘, 7, 6.8}] (* y(2005)=18500 KW *)
{18.46}
model1[[1]][{18.46328449471389‘, 6.4, 7}] (* y(2006)=30500 KW *)
{30.47}
model1[[1]][{30.465007408843576‘, 6.5, 6.4}] (* y(2007)=83900 KW *)
{83.72}
model1[[1]][{83.71931076357305‘, 6.0, 6.5}] (* y(2008)=295000 KW *)
{294.43}
model1[[1]][{294.4250328435696‘, 4.5, 6.0}] (* y(2009)=656800 KW *)
{656.38}
model1[[1]][{656.3822847076975‘, 4.5, 4.5}] (* y(2010)=? KW *)
{735.63}
```

Table 2 summarizes the trend of empirical and theoretical data computed by the neural network (data expressed in MW). Data refer to years ranging from 2002 to 2009, with a forecast for 2010 of 735.63 MW. The square index of deviation is now $I_2 = 0.002074$, better than that computed through the analytical model in Section 4.

Table 2: Trend of empirical and theoretical data computed by the neural network with two hidden layers and RMSE=0.233946

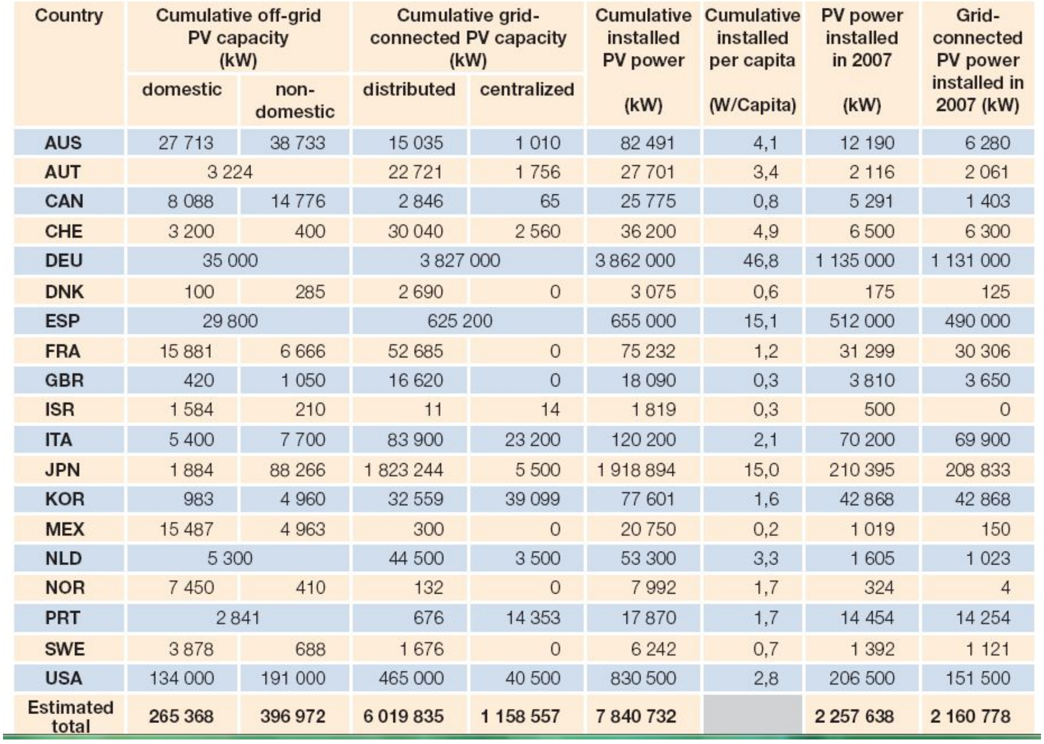
YEAR	2002	2003	2004	2005	2006	2007	2008	2009
Empirical	3.62	7.60	12.00	18.50	30.50	83.90	295.00	656.80
Theretical	3.80	7.34	12.14	18.46	30.47	83.72	294.43	656.38

7 Conclusions and Future Work

As seen in the presentation neural networks are a powerful and versatile tool for forecasting. In activities that require the use of predictive models neural network-based applications are increasingly important, and the Mathematica environment offers a wide variety and wealth of research tools and modeling.

From the analysis of the two indices of square deviation (see Section 4 and subsection 6.7) it can be seen that the neural network model, although working with a small number of input data, is an order of magnitude more efficient than the discrete mathematical model.

This conclusion stimulates our work to evolve not only in the study of photovoltaics in Italy but also in European Community countries for which data were published by IEA [3] (see Figure 14).



Country	Cumulative off-grid PV capacity (kW)		Cumulative grid-connected PV capacity (kW)		Cumulative installed PV power (kW)	Cumulative installed per capita (W/Capita)	PV power installed in 2007 (kW)	Grid-connected PV power installed in 2007 (kW)
	domestic	non-domestic	distributed	centralized				
AUS	27 713	38 733	15 035	1 010	82 491	4,1	12 190	6 280
AUT	3 224		22 721	1 756	27 701	3,4	2 116	2 061
CAN	8 088	14 776	2 846	65	25 775	0,8	5 291	1 403
CHE	3 200	400	30 040	2 560	36 200	4,9	6 500	6 300
DEU	35 000		3 827 000		3 862 000	46,8	1 135 000	1 131 000
DNK	100	285	2 690	0	3 075	0,6	175	125
ESP	29 800		625 200		655 000	15,1	512 000	490 000
FRA	15 881	6 666	52 685	0	75 232	1,2	31 299	30 306
GBR	420	1 050	16 620	0	18 090	0,3	3 810	3 650
ISR	1 584	210	11	14	1 819	0,3	500	0
ITA	5 400	7 700	83 900	23 200	120 200	2,1	70 200	69 900
JPN	1 884	88 266	1 823 244	5 500	1 918 894	15,0	210 395	208 833
KOR	983	4 960	32 559	39 099	77 601	1,6	42 868	42 868
MEX	15 487	4 963	300	0	20 750	0,2	1 019	150
NLD	5 300		44 500	3 500	53 300	3,3	1 605	1 023
NOR	7 450	410	132	0	7 992	1,7	324	4
PRT	2 841		676	14 353	17 870	1,7	14 454	14 254
SWE	3 878	688	1 676	0	6 242	0,7	1 392	1 121
USA	134 000	191 000	465 000	40 500	830 500	2,8	206 500	151 500
Estimated total	**265 368**	**396 972**	**6 019 835**	**1 158 557**	**7 840 732**		**2 257 638**	**2 160 778**

Figure 14: IEA data for European countries.

The research here undertaken with the study of neural networks applied to photovoltaics is not an end in itself but it goes in one more comprehensive direction.

In fact, among the new research challenges facing us are the applications of neural networks in the following areas of development:

- provide, as a function of some initial parameters, the trend in the price of electricity in a free market of this commodity;
- to ascertain whether there may be conditions allowing forecasting electricity consumption and thus create an application that (based on a neural network model) can predict the possibility of micro and macro black-out through the formalization of an index able to express the "Critical Energy Factor".

References

[1] Denholm, P. Margolis, R.M. Milford, J. 2009. Quantifying avoided fuel use and emissions from solar photovoltaic generation. *Environ Sci Technol* (43) 226–232.

[2] Solarbuzz. 2009. *Solar electricity price index.* Link: `http://solarbuzz.com/SolarPrices.htm` (link accessed on January 27, 2009).

[3] IEA. International Energy Agency. Link: `http://www.iea.org/` (link accessed on February 3, 2011).