Dr. Marina Resta,
c/o Professor Davide Sciutti
Istituto di Matematica Finanziaria,
Facoltà di Economia,
via Bertani 1,
16121
Genova,
ITALIA
resta@economia.unige.it

SELF ORGANIZING EVOLUTIONARY MODELS IN FINANCIAL MARKETS FORECASTING.

Abstract. A Self-Organizing Feature Map (SOFM) is a kind of neural architecture making an acceptable compromise between information complexity and the need of explanatory capabilities for an assigned system.
Such properties make SOFM a powerful support to be used over specific contexts (financial markets...) where mathematical modeling efforts have often been vanished, and no alternative choices are offered to researchers, but for technical tools, the orthodoxy of which is still far to be completely accepted.
SOFM most relevant property relies, perhaps, in learning procedure flexibility. On the other hand, we have to face against such a structural rigidity affecting the map, because its planary dimensions are arbitrarily settled during earlier steps of the process. This evident bottleneck of the procedure led us to develop an algorithm, which overcomes those limits by combining the approach of SOFM to that proper of Evolutional Models.
The result is a process integrating knowledge from Self-Organization Theory and Genetics (ISOG), the efficiency of which has been tested on financial markets.

I. Introduction.

Self-Organizing Feature Maps (SOFMs) join biological coherency, together with the power of a classifier, thanks to which comprehension of complex systems can be relevantly simplified.
A SOFM is made up by elementary units (i.e. neurons). Each neuron, in turn, is a string having the same length of inputs in the dataset, and its elements are initialized at random values.
This is a crucial procedure to find out important blocks of information. We can think each component of neurons to act as a marker, i.e. it runs as a fenotipically relevant gene: we know its position, and we can use it as lighting guide to map new mutations.
The number of neurons (hence the dimensions of the map) assures us to manage with molecular polymorphism, independently from the position of the element. It is then possible to fit correctly each input. This result is not trivial. We must remember that arranging 100 input samples in 1×5 vectors means to deal with over 75 million alternative features for each neuron: very similar inputs shall be distinguished on condition it exist a tool able to face with such great deal of information.
SOFMs win information complexity thanks to a flexible learning procedure, during which the amplitude of neuron neighbourhood decreases step by step. On the other hand, the process finds a bottleneck in its structural rigidity: map dimensions are arbitrarily defined in the earlier steps of the procedure, whilst the rightness of this choice can be judged only once the learning has been ended.
However, SOFMs are powerful and easy to use tools. They are employed in very different fields (Digital Imaging, Signal Processing, Markets Forecasting).
In particular, their use in specific economical contexts (exchanges market, market indexes...) revealed SOFM capabilities in developing trading systems, even if they provide qualitative information which can only integrate the set of knowledge we need to make trading strategies.
Our goal is that of studying the possibility of combining the approach of SOFM to have a tool which could offer both quantitative and qualitative information about financial markets.
We have obtained satisfying results by developing a procedure, which combines SOFMs learning procedure engine in an evolutional model.
This work is organized as follows. Section II depicts the algorithm we have developed. In section III we compare both the results we have joined by using a traditional SOFM, and those obtained by using the new procedure. Networks performances will be evaluated by referring to financial criteria. Finally, in section IV we'll give conclusive remarks.

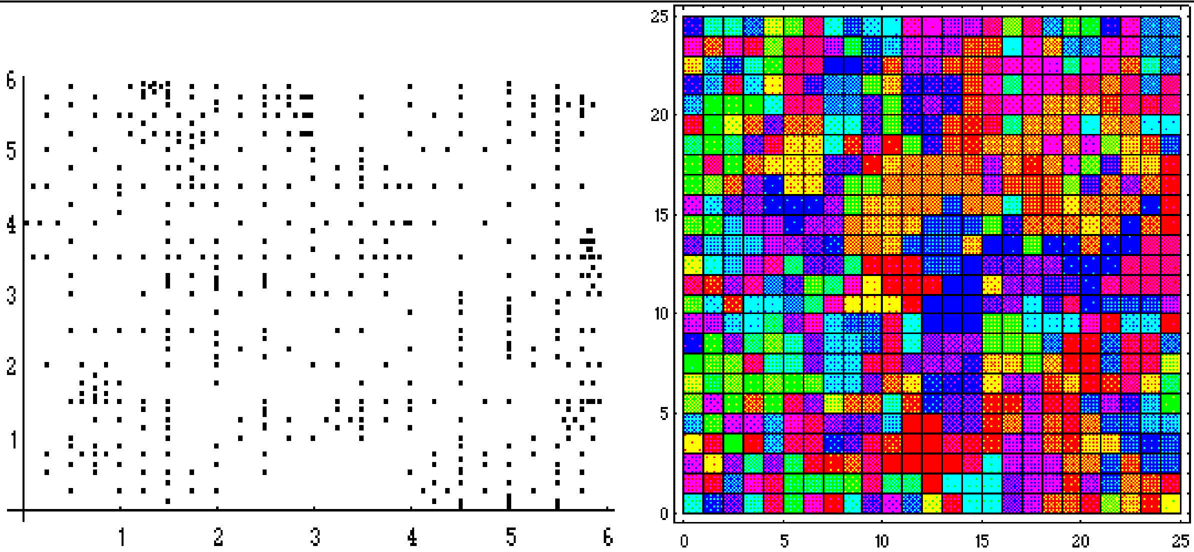
II. Genetic Algorithms and Self-Organization: an integrated process (ISOG).

Our aim has been that of combining in the same procedure winning properties of two different classifying systems: SOFM and Genetics Algorithms.
We have considered a set of input, i.e. a data set arranged in vectorial way, starting from which the classifier could learn proper solutions.
As object of the evolutive process we have chosen a " virtual community" of neurons dimensionally compatible to inputs data, and we have initialized them to random values.
The system maintains properties of SOFMs, except for a better structural flexibility. The structure is initially made up by a kernel of neurons arranged into a planar lattice of reduced dimensions, and its growth depends from evolutionary laws, as well as in a biological system.
In this way we don't fix final map dimensions which follow as natural evidence of the learning procedure. The process is organized in many steps.
Step 1. An input sample is presented to the net.
Step 2. Each neuron is evaluated according to SOFM similarity criteria.
Step 3. The procedure finds a winner neuron, i.e. the neuron offering the best approximation to input.
Step 4. New neurons are originated from the leader of Step 3., and added to neural community. The number of neurons belonging to new generation depends from a growth factor, which is settled at the beginning of the procedure, as well as their spatial position in the map. Chromosomic features of new neurons are in part hereditarily transmitted by the originary cell (winner neuron), and partly obtained from a list of random values within a prefixed range: the procedure is similar to biological crossing over.
Step 5. A number of neurons is deleted to making space for new member of the community. A control procedure operates, valuing the fitness of each neuron, and canceling those appeared as leader a prefixed number of times.
Step 6. The process starts again from step 1, until all inputs have been shown to the map.

III Simulation Results.

Our simulation has been joined on MIB daily fluctuations, from March 1st 1990 up to May 22nd 1996. We used earlier 500 samples as learning set, while remaining values have been arranged into groups we have shown to the net 100 per time, as control set.
A 5 days long window of observation has been considered.
We have made our tests by running both traditional (SOFM) and new algorithm (ISOG). In the first case we have referred to a 25×25 map, in the latter case we have referred to a kernel net made up by 25 neurons. Final community has been made up by 561 elements.

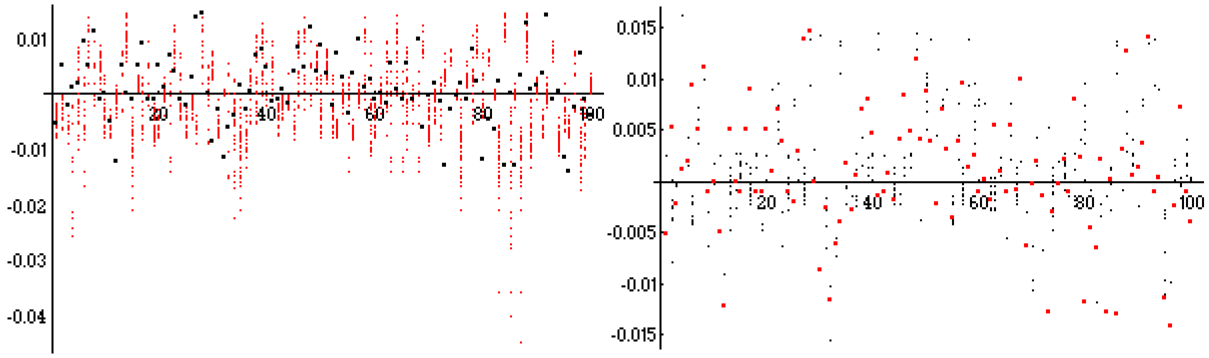
From a geometrical point of view, different characteristics of ISOG and SOFM are shown in Picture 1.



a) ISOG b) SOFM

Picture 1. Geometrical depicts of the two maps.

Different colors indicate different inputs fitted by the map, according to the principle of topological invariance, which is proper to SOFM: similar inputs are mapped onto neighboring areas of neural space, i.e. areas which appear to have similar color gradation refer to similar inputs (Picture 1 . b.).
ISOG doesn't maintain the idea of topological lattice, but we gain information about the number of neurons fitting each input.
We now can count on molecular polymorphism (the agglomerate of points in Picture 1. a.), and hence on better probabilities to find the right feature for each input from sensory space.
Under this point of view, then differences are less relevant than what we could think.
Picture 2 shows the position of input respect of that of neurons (leader and neighborhood) giving better approximation.



a) ISOG b) SOFM

Picture 2. Input and neurons fitting it. The input is prepresented as a point of greater dimensions than neurons.

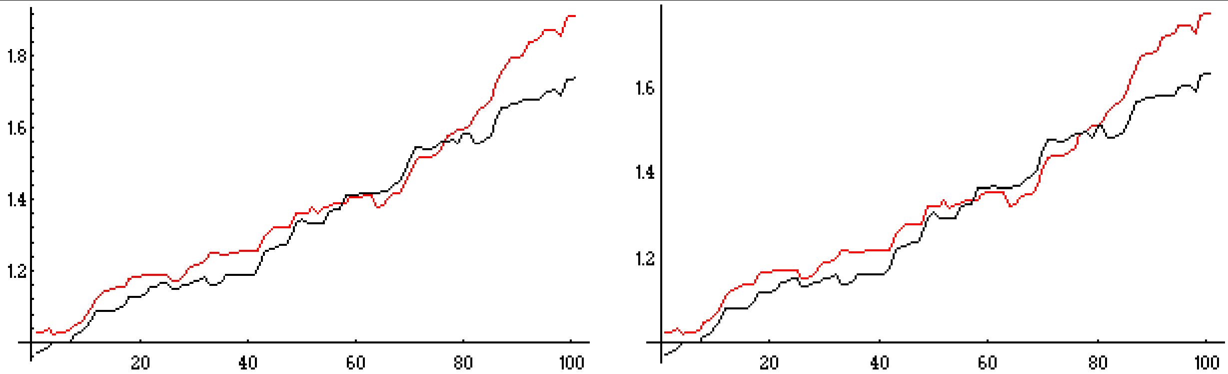
Both the simulations offer satisfying results, by fitting well input data set, but most relevant difference is in the number of neurons joining each input : in the ISOG -based net the neural community for each input is significantly greater than in the SOFM.
This could mean ISOG to have a better vocation to give both quantitative and qualitative results for financial markets.
Finally, we have tested control sets, evaluating the results only by means of financial criteria.
We have supposed to employ all information given by the nets to act selling / buying strategies over our market.
We have also assumed a number of hypotheses :

- purchasing and selling tasks have been considered

a) costless ;
b) with daily transaction costs of 0.001 lit for each liras of capital employed ;

- each time the net told us to buy we bought, each time the net told us to sell we sold. We skip current position every time the net, coming from uncertainty in sensory space, provided a standby signal ;
- we kept zero interest rate in standby positions, paying right responses at the current daily fluctuation rate, and punishing any loss at the same (negative) rate, under the assumption of composed interest rate rule.

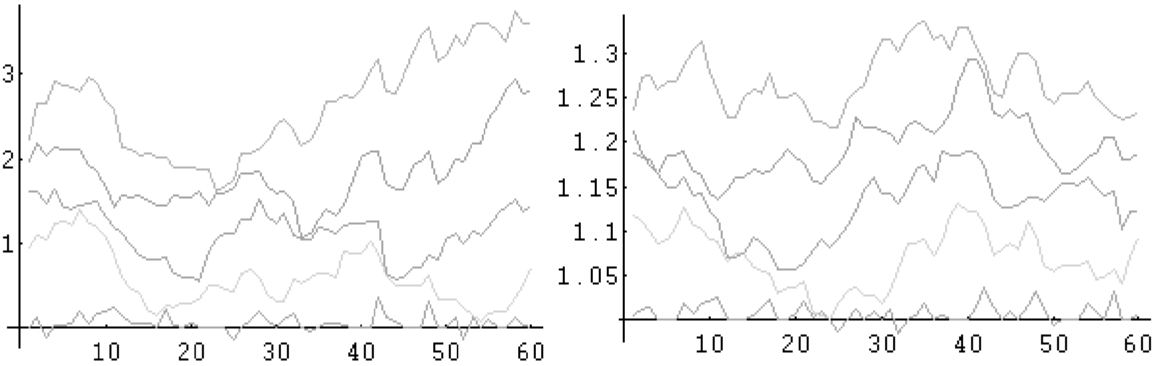
Picture 3 shows the results under the assumption of costless tasks (a) and the presence of transaction costs (b).



a) No transaction costs b) Transaction costs have been incorporated in the model

Picture 3. Comparative results. ISOG performances (red) vs. SOFM (black) over 100 days. In the picture results are shown by assuming different hypothesis on transaction costs.

These results are stable, i.e. there isn't any dependence from the starting day of the simulation, as picture 4 clearly shows.



a) ISOG b)SOFM

Picture 4. Networks performances are independent from starting point of the simulation. That condition is explained on a 60-day wide range. Every day from 1st to 60th (X axes) can be chosen to enter the market (and staying on it up to a maximum of 40 days), the net is able to ensure complete stability of the results you can get by trading. These performances have been tested by monitoring capital employment on the market periodically: the curves have been obtained by checking the market every 10 days on a 40 days long period.

We got satisfying performances from both kind of neural architectures, even if ISOG more powerful fine tuning capabilities influe on the results in the last part of the sample set. ISOG extreme specializing vocation can also be negatively evaluated by observing results among 40th and 60 th days, where SOFM appear to do better results.
ISOG overall performances however appear to be more significant than those of SOFM are.

IV Conclusive remarks.

We have tried to suggest an alternative for studying and forecasting on financial markets. Our aim is offering a predictional tools which can act both from a quantitative, and a qualitative point of view.
The idea is that of combining winning properties of SOFM and Genetic Algorithms. Efforts in this direction have produced ISOG (Integrated Self-Organizing and Genetics) algorithm.
We valued the new algorithm only on the base of financial criteria, finding its performances better than those of a traditional SOFM did. Only experimental proofs and control over the hypothesis behind shall support or contrast those affirmations.

References

1. BLACK F. - SCHOLES M. , *The pricing of options and corporate liabilities,* Journal of Political Economy, vol. 81, pp 637-659, 1973.
2. DAVIDSON C. , *Views from the Frontier: Commentary on the New World of Forecasting and Risk Management* , May 1996, abstract from the following URL : http://www.olsen/library/prof/ (Olsen & Associates web site).
3. KANGAS J. A.- KOHONEN T.K.- LAAKSONEN J.T., *Variants of self- organizing maps,* IEEE Transactions on Neural Networks vol. 1, n° 1, pp 93- 99, March 1990
4. PRUITT S. W. - WHITE R. E. , *The CRISMA trading system : who says technical analysis can't beat the market?*, The Journal of Portfolio Management, pp 55 -58, Spring 1988.
5. RESTA M., Un sistema di trading neural network oriented per le decisioni di borsa, Bozzi Editore, Genova, 1997.