GIFINT

Work Package WP100

ANN-prediction of geomagnetic indices

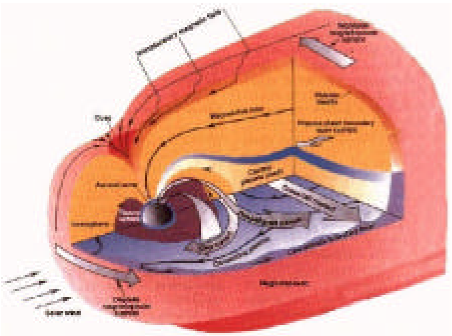
G. Consolini, E. Amata, M.F. Marcucci,
G. Pallocchia, I. Bertello, M. Candidi

Istituto di Fisica dello Spazio Interplanetario
Consiglio Nazionale delle Ricerche
Roma, Italia

Outline

- The physical scenario
- Previous results on the prediction of the geomagnetic indices
- The ANN architecture
- Some preliminary results
- Conclusions

The physical scenario



Taken from NASA ISTP site: http://www-istp.gsfs.nasa.gov

Some previous results

In the past different approaches have been applied to study the solar wind-magnetosphere-ionosphere (SWMI) coupling:

- statistical correlative analyses (*Baker*, 1986)
- linear filtering (*McPherron et al.*, 1988)
- nonlinear filtering (*Klimas et al.*, 1992; *Goertz et al.*, 1993)
- artificial neural networks - ANN (*Lundstedt*, 1992; *Wu & Lundstedt*, 1997; *Takalo & Timonen*, 1997; *Gleisner & Lundstedt*, 2001; *Lundstedt et al.*, 2002).

Most of the previous work was devoted to the prediction of geomagnetic indices (i.e. AE and DST) as a function of solar wind parameters.

Our task

We intend to build a service based on ANN to forecast Dst and AE based on ACE Solar Wind data.

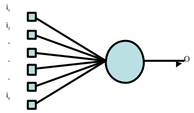
In this presentation we will concentrate on the Dst index prediction.

The ANN concept

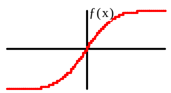
The ANN approach seeks 'merely' a simple improvement rule whose application leads to action rules which are appropriate under the conditions experienced.

An ANN may be considered as a "black-box", which is able to estimate the answer function of a given system without knowing the detailed processes which govern it.

The elementary unit of an ANN is the neuron (or node), which responds to the external stimuli via a specific transfer function (generally a hyperbolic tangent transfer function).



$$O(t+1) = f\left[\sum_k w_k i_k(t)\right]$$



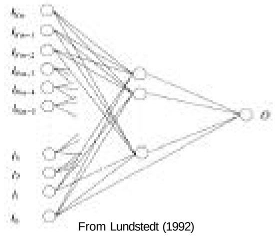
A schematic picture of a simple neuron

ANN architectures

In the past studies, several different ANN architectures were used to try to deduce geomagnetic indices (AE, DST) from solar wind parameters.

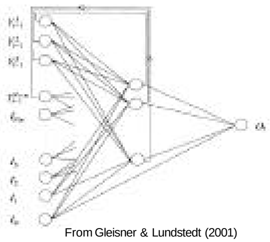
Among the more commonly used we remind the time delay network (i.e. a standard perceptron with input parameters at times t_0, t_1, ...t_n), and the Elman recurrent network.

A three layer time delay network - TDN



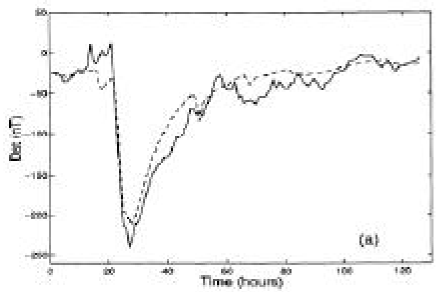
From Lundstedt (1992)

An Elman recurrent network - ERN



From Gleisner & Lundstedt (2001)

On the basis of previous works, good predictions are attained for the DST index (Wu & Lundstedt, 1997).

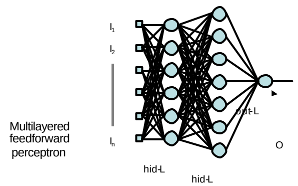


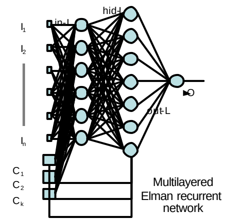
2-hours ahead prediction from Wu & Lundstedt (1997)

The work we did until now.

We used various ANN architectures:

- a simple multilayered feed-forward perceptron;
- a time delay network;
- an Elman recurrent network;
- a modified Elman recurrent network.





The network training.

The training was performed minimizing the cost function by adjusting the weight coefficients of the linear output transfer function:

$$E = \frac{1}{2}\sum_i \sum_m \left(T_i^m - O_i^m\right)^2$$

The minimizing technique was the gradient method.

Different methods for selecting the training set were used.

The various tests performed

In our simulations we made several tests as regards:

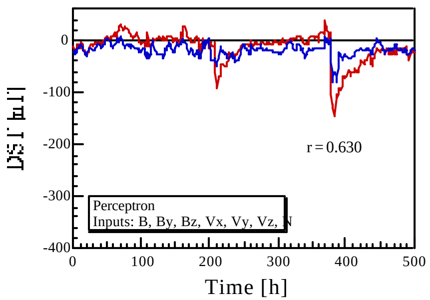
- the input variable set;
- the number of neurons in the hidden layers;
- the μ parameter, i.e. the inertia of the weight coefficients as a function of learning step;
- the learning rate parameter η;

The input data.

As input variables we used WIND and ACE magnetic and plasma parameters from the OMNI dataset.

Different methods for selecting the training set were used.

An example from one of our first tests.

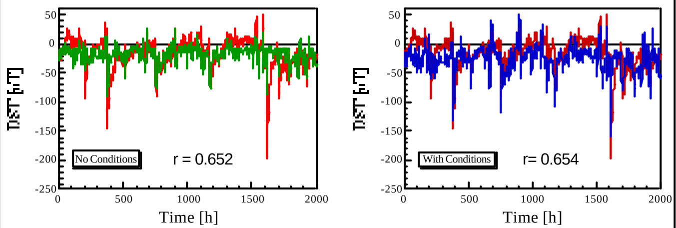


Training set: 100 patterns of 2000 pnts sorted over 20000.

The importance of the training set selection.

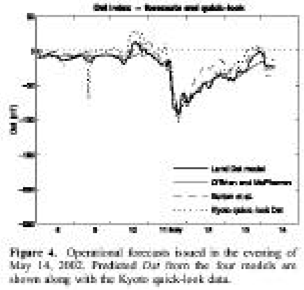
Training set pertaining to the right hand plot: 100 patterns, 1200 points each, sorted over 20000 under the following condition:

$DST(i+5) > 13.5$ and $DST(i-15) < -58.05$



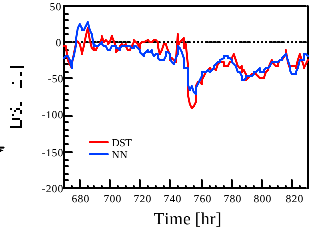
Input parameters: B, By, Bz, N with 20 input layers [I(i), I(i-1), I(i-2), I(i-3) and I(i-4)]

A comparison between our results and a case studied in literature



A case study from Lundstedt et al., 2002

A period with similar characteristics from one of our tests.



Conclusions

- High sensitivity of the forecast precision on the training set selection and training condition.
- Inability at the present moment to have an architecture able to forecast small and large fluctuations at the same time.
- Our forecasts seem to give reasonable results, at least as regards the goals of the WP.
- A prototype of the GIFINT web site showing the real time prediction is already operational. The provisional URL is:

http://pcfede.ifsi.rm.it

Next step.

Extension to the AE index prediction.