Geophysical Research Abstracts, Vol. 9, 07046, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-07046

Input selection based on mutual information for solar activity prediction

M. M. Rezaei Yousefi (1), A. H. Vahabie (1), M. Falahi (1), C. Lucas (1,2) and B. Nadjar Araabi (1,2)

(1) Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran, (2) School of Intelligent Systems, Institute for studies in Theoretical Physics and Mathematics (IPM), Tehran, Iran (m.rezaei@ece.ut.ac.ir, h.vahabie@ece.ut.ac.ir, m.falahi@ece.ut.ac.ir, lucas@ipm.ir, araabi@ut.ac.ir / Fax: +98 21-88633029 / Phone: +98 21-61114313)

Prediction of space weather is very useful to prevent damages on satellites, communication and electric power systems. Purposes of many recent researches are better analysis and more accurate prediction of solar activity. Foundation of prediction is on modeling problem. Choosing a set of most relevant and non-redundant input variables is necessary to build an appropriate model with high generalization performance, and to improve the interpretability of the selected inputs set. When the choice of input variables is confined to the lagged values of the process to be predicted, a nonlinear analysis of the most significant factors is crucial for improving the prediction quality. In this paper, we use information theoretic criterion as a nonlinear analysis to select a subset of input variables that have the richest information about the output where methods based on linear relations (such as correlation analysis) may give misleading information. Satisfactory results of sunspot numbers prediction confirm our claims. It is noticeable for us that eleventh lag of sunspot numbers time series is selected as an important input variable. Results of this method are in accordance with physicists explanation on existence of approximately 11-years solar activity cycle.