Information Theoretical Approach to Attention in Geomagnetic Storm Forecasting: a Tutorial

A. Arami, M. Rezaei Yousefi, B. S. Kasmaei, M. Mirmomeni, Caro Lucas

Control and Intelligent Processing Center of Excellence, ECE School, University of Tehran

Outline

- The necessity of attention control
- Motivation
- Active perception and attention
- Attention control formulation
- Data driven approaches
- Space weather-Geomagnetic storm
- Problem definition
- Simulation results
- Conclusion

The necessity of attention control

Complexity of nature

- Complexity because of unlimited information in nature
- We feel the complexity of the nature because of our powerful sensors.
- We are capable to detect high dimensional data, using our sensors
- Is the world very complex from a blind primitive insect point of view?

The necessity of attention control

Limited power of computation

- What can we do about huge collected data by our sensors, while the brain power of process is limited?

- We can turn off some of our sensors, on the other hand we neglect them and focus on the other ones.
- Active perception of agent forms when it learns how to use its sensors actively in a complex environment.

Primitive form of Active Perception

Dumb Cricket? Smart Cricket? *(Clark, 2001)*

Fact:

Female crickets (somehow) "recognize" the chirping songs of males in their species and move toward them. They do not move toward males of other species.

Classic Symbolic AI explanation:

Female cricket must:

1. Hear songs and classify them as *same*- or *different*-species.
2. Determine the location of the conspecific male.
3. Move toward that male.
4. Maybe even **plan** a route to that male before moving.

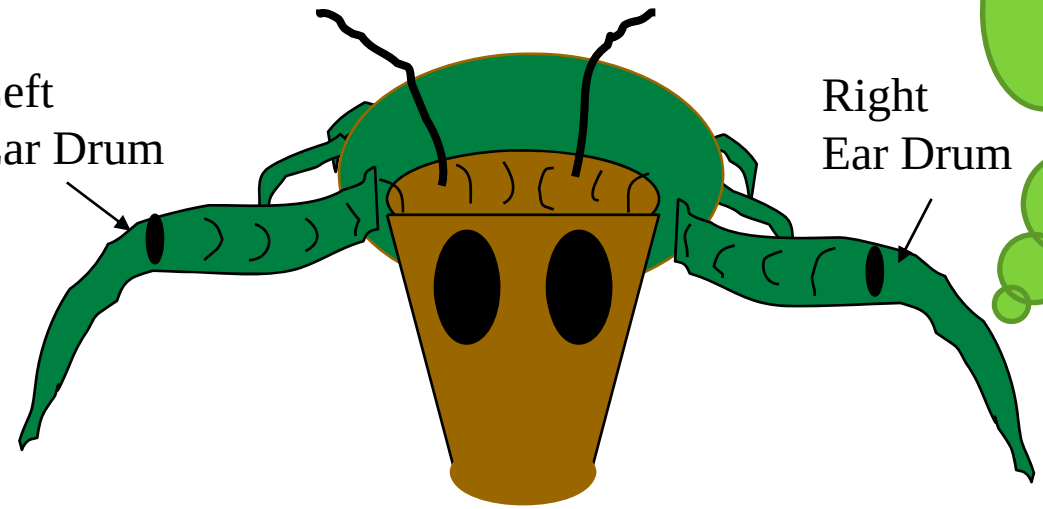
Situated and Embodied Intelligence explanation = Biological explanation

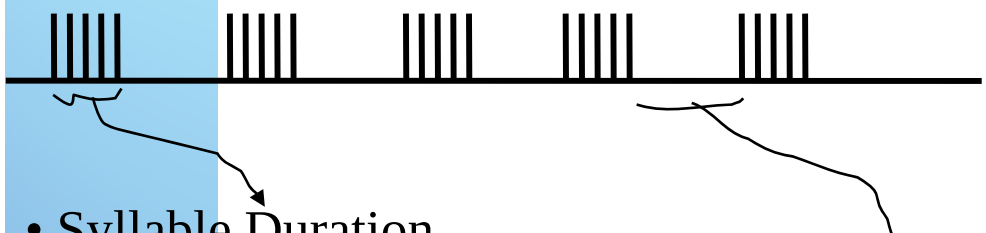
The physical interactions between the song (environment), auditory canal (body) and relatively simple neural circuitry (brain) of the cricket lead to approach behavior in only those special cases where the song of the male is compatible with the body and brain of the female.

Cricket Phonotaxis

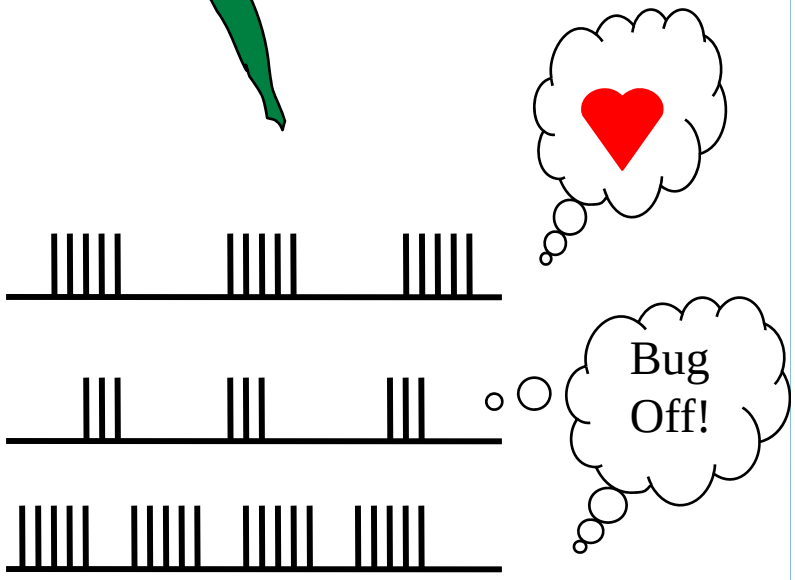
- Webb, B. (2001). Biorobotics: Methods & Applications
- Female Crickets only respond to songs with partic
  frequencies and **syllable durations**.

By moving elbows, and traveling in environment it works like an active sensor and eliminate the need for complex processing





- Syllable Duration
- Carrying Frequency = 1/Inter-syllable period



Motivations

- Huge and complex collected data
- Different constraints on sensing and processing
- The cost of processing
- Can we process only the limited relevant parts of data in order to reduce costs?

About Active Perception and Attention

Perceptual activity is exploratory, probing, searching and also related to prediction of events.

percepts do not simply fall onto sensors as rain falls onto ground.

We do not just see, we look.

What to look? and what to do?

learn the attention control.

learn the action and policy to achieve

the perception and reward

Attention:

-attention could be described as a filter

-the output of attention filter could be easier to processing

About Active Perception and Attention

Attention:

- top-down (Goal directed- need to be learned)
- bottom-up (Context based; need a priori knowledge)
- combination of top-down and bottom-up

We formulate the attention control problem as an optimization problem and the main goal is to achieve maximum expected reward, on the other hand, our aim is to improve the performance of system.

About Active Perception and Attention

This is a hard problem, because in general the rewards are the results of physical action and it's effect on environment, but attention is a perceptual activity and there is no necessary direct reward and punishment for it.

Attention Control Formulation

formulation

$$Subspace^{*} = \arg \max_{Subspace} (relevant\ information\ of\ subspace);$$

$$S.to: Computational\ Constrains\ and\ Sensory\ Constrains$$

→

$$Subspace^{*} = \arg \max_{Subspace} ([relevant\ information\ of\ subspace] - \gamma[Costs])$$

- How we can evaluate the relevancy of information of a subspace of data with respect to the goals?

1) Using some indices (such as Corr. , MI) which are estimators of level of information relevancy

2) Interactive learning

3) Global search algorithms

In 2 and 3, we need to evaluate the learning or find points in search space by closing the loop of estimation, prediction etc.

MI *vs.* Correlation

- Correlation is one of the criterion which indicate the linear relation between variables.
- Mutual information (MI), as a generalization of correlation [Schwarz et al., 1994], is one of the most powerful and suitable criterion which can indicate any kind of relation, linear or nonlinear, between variables.
- Encountering a nonlinear system with complicated multi-dimensional dynamics, like solar-terrestrial system, requires MI to recognize the existence of such nonlinear interrelations that is superior to correlation analysis.

MI *vs.* Correlation

From the other standpoint:

- The calculation of MI is extremely cumbersome, and leads to huge computation costs
- Although the other methods used for estimation of MI decrease the computational cost, in comparison with Correlation computation costs they are massive

as an instance in a simulation:

Method	correlation-based	MI-Kraskov
CPU time	1 sec	38 sec

→ A tradeoff is needed between use of MI and Corr.

Formulation of attention control using MI, an estimation of MI or Correlation

MI-Attention:

$$\widehat{Subspace}^* = \arg\max_{Subspace} \left([MI \text{ between subspace and output event}] - \gamma[Costs]\right)$$

MI_estimate-Attention:

$$\widehat{Subspace}^*_{approx2} = \arg\max_{Subspace} \left([MI \text{ between subspace and output event}] - \gamma[Costs]\right)$$

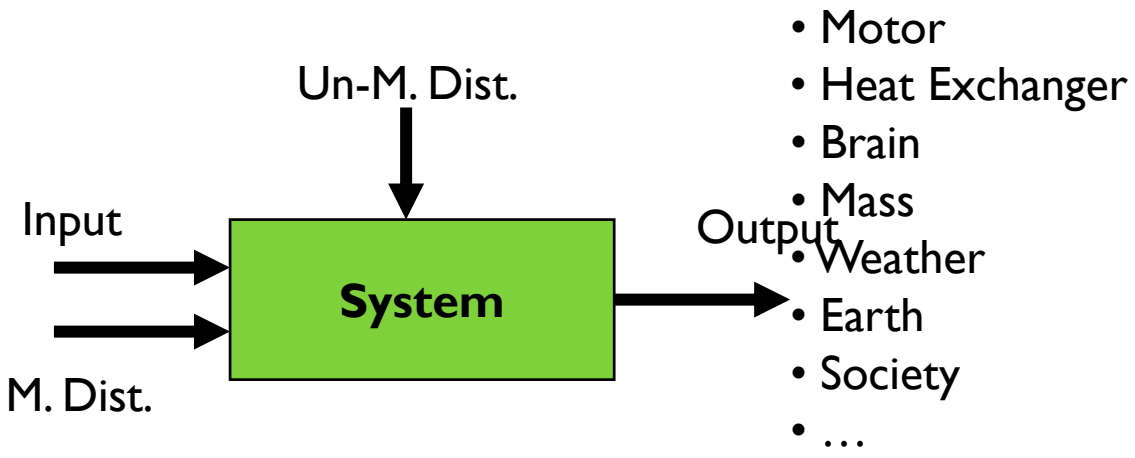
Corr-Attention

$$\widehat{Subspace}^*_{approx1} = \arg\max_{Subspace} \left([Corr \text{ between subspace and output event}] - \gamma[Costs]\right)$$

Data Driven Approaches

Inferring models from observations and studying their properties is really what science are about (Ljung).

A system is an object in which variables of different kinds interact and produce observable signals.

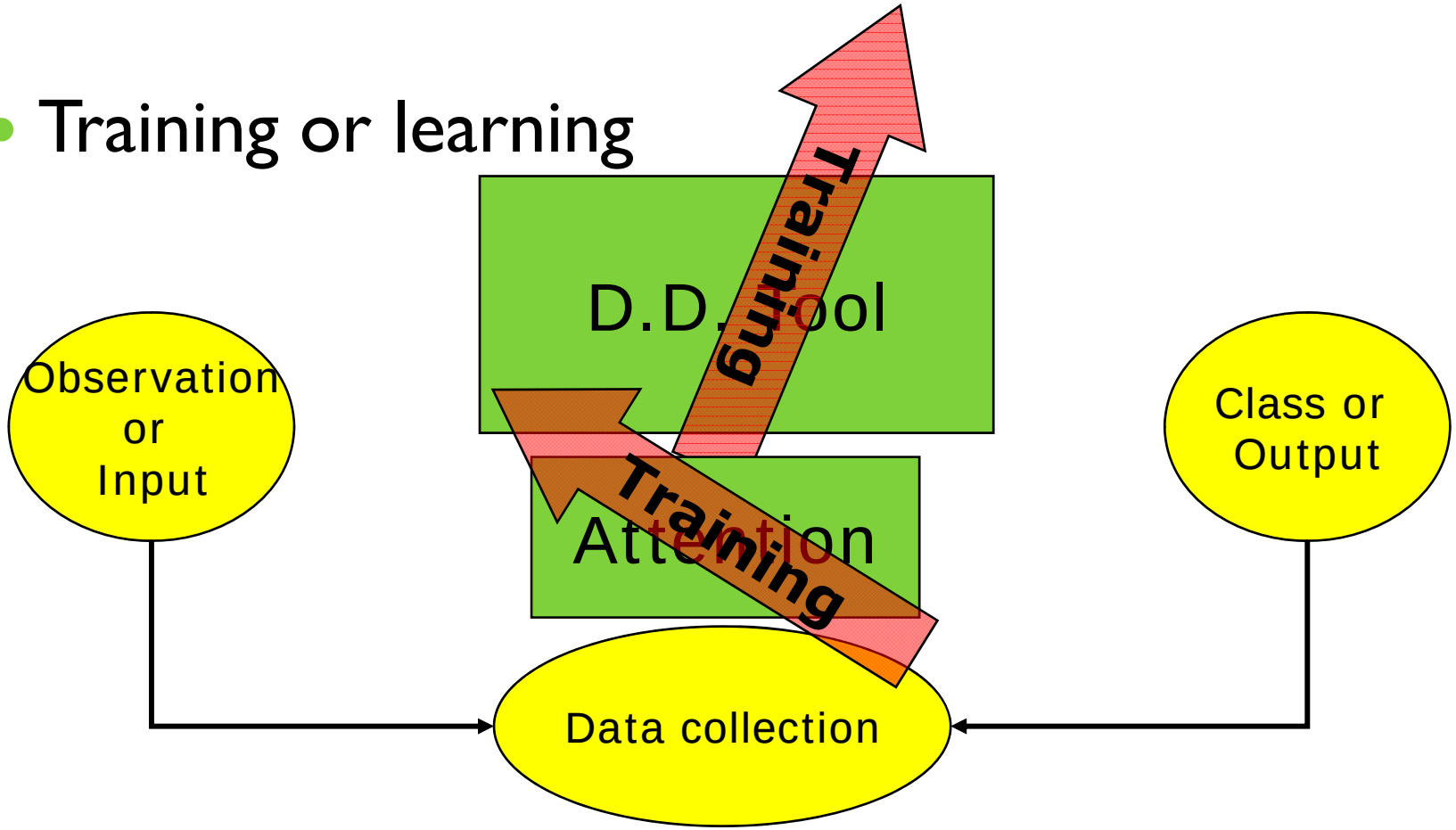


Data Driven Approaches

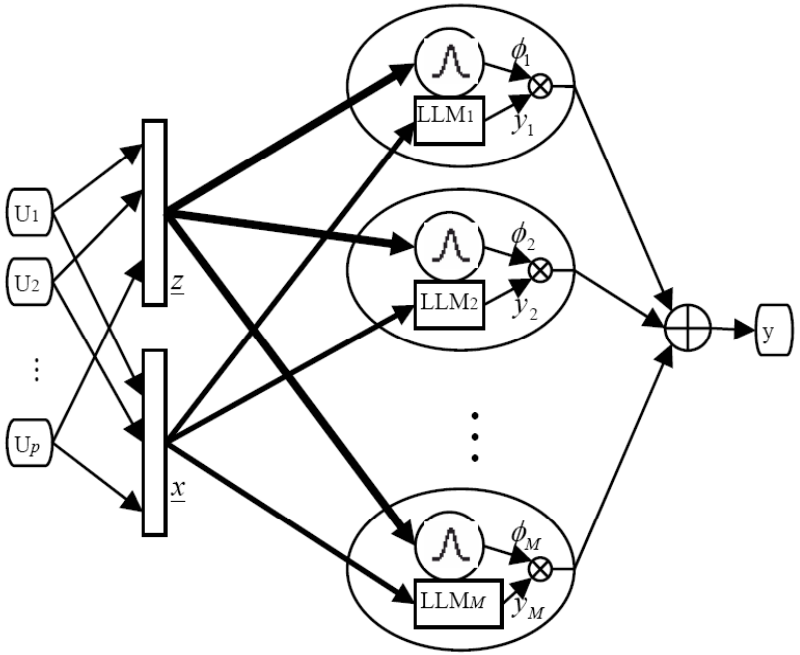
- White Box Modeling: Only Considering Physical Rules
- Gray Box Modeling: Considering Data and physical rules simultaneously
- Black Box Modeling: Considering Only Data

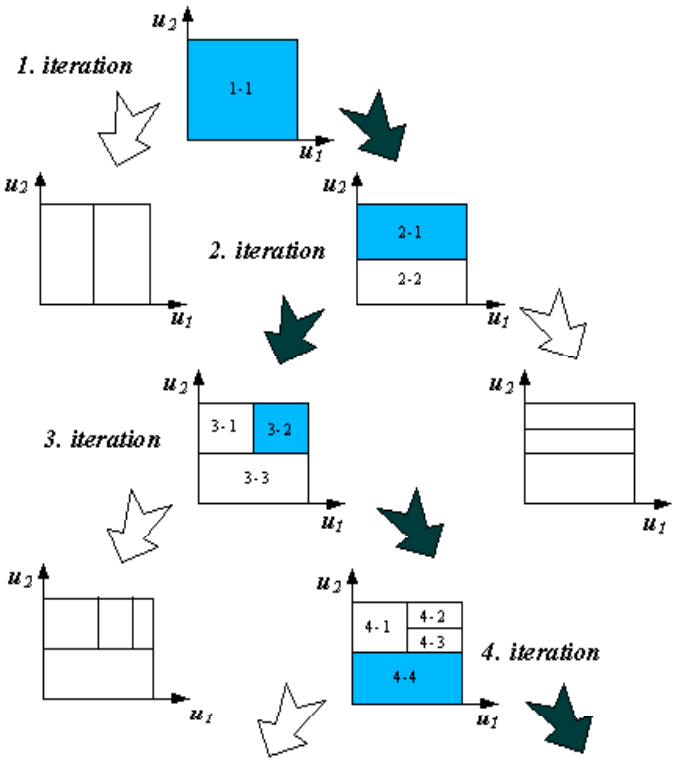
Training in Data Driven Approaches

- Training or learning



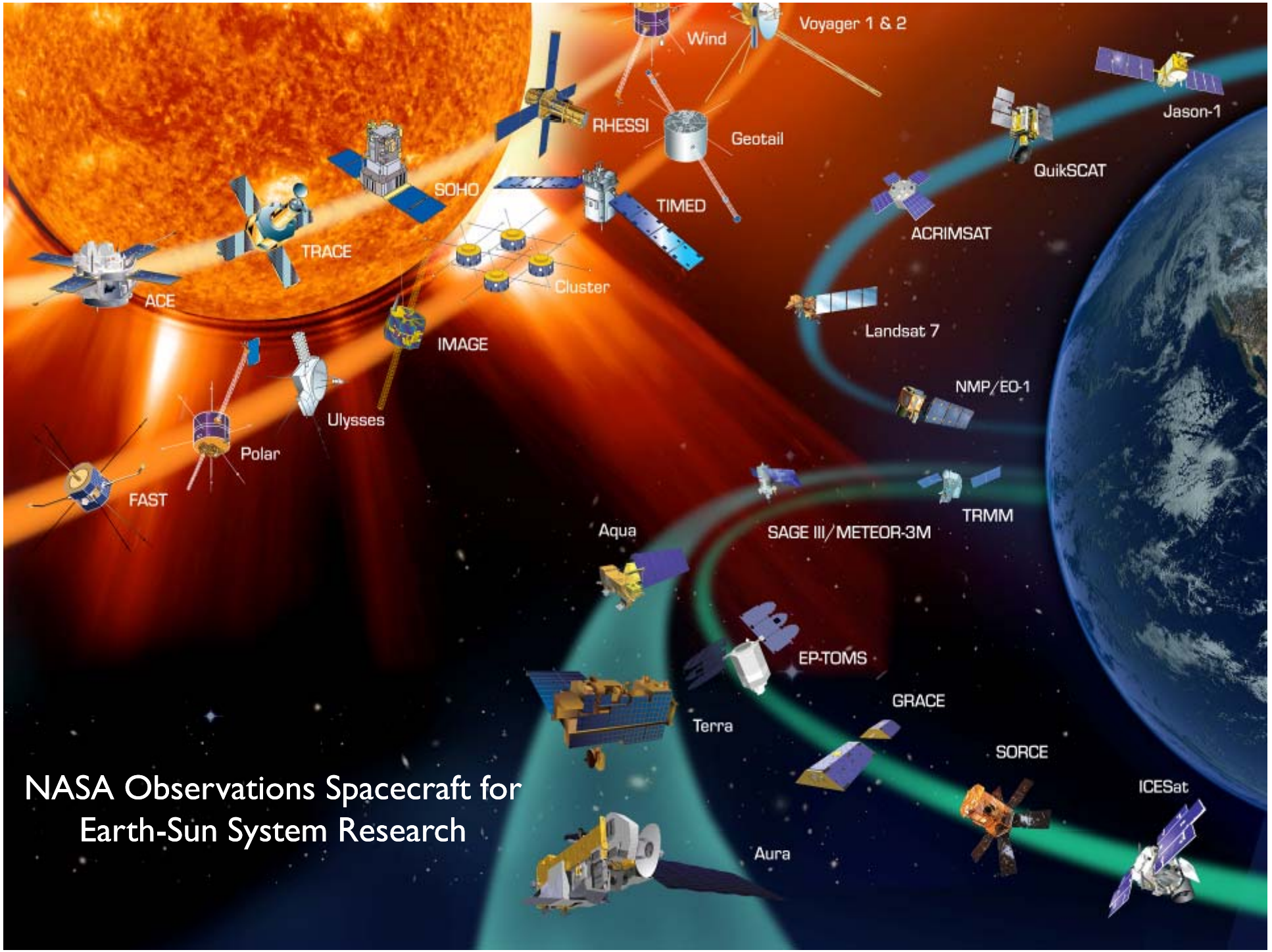
Locally Linear Neuro Fuzzy (LLNF) Models with LoLiMoT Learning Alg.





Space weather-Geomagnetic storm

- Since our life and technological systems such as satellites, communications and power transmission are directly affected by the conditions of space weather, we need to increase our knowledge about the possible hazards and the Sun as the their main source. Satellites and ground-based stations have recorded various types of data from the solar-terrestrial system during recent decades.
- The Magnetosphere is a shield for Earth in front of solar wind and high energy particles. Weakening of geomagnetic field caused by solar wind pressure on magnetosheath is called geomagnetic storm which plays an important role in changing upper Earth environment such as Ionosphere which might be dangerous to the human life.


Wind
Voyager 1 & 2
RHESSI
Geotail
Jason-1
QuikSCAT
SOHO
TIMED
ACRIMSAT
TRACE
Cluster
ACE
IMAGE
Landsat 7
NMP/EO-1
Ulysses
Polar
FAST
TRMM
SAGE III/METEOR-3M
Aqua
EP-TOMS
GRACE
Terra
SORCE
ICESat
Aura
NASA Observations Spacecraft for
Earth-Sun System Research

Space Climate/Weather

Space Climate/Weather refers to changes in the space environment and effects that those changes have on Earth and mankind's activities.

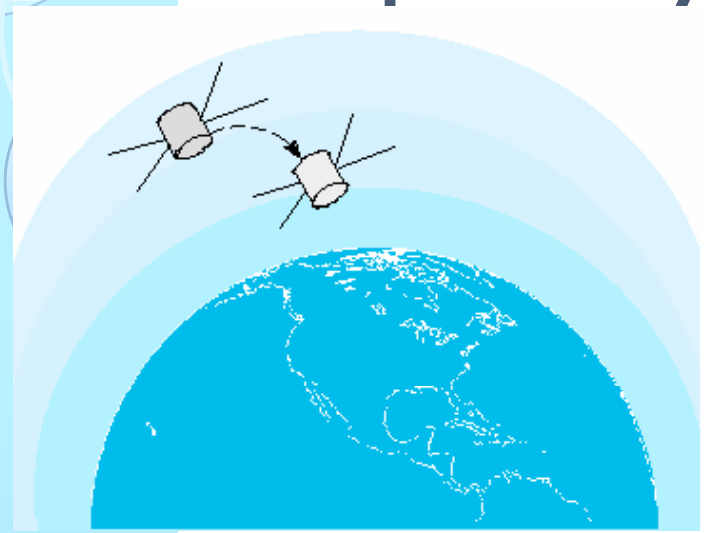
These affect **Earth climate on various temporal and spatial scale** as well as **communications, navigation and many other space and ground based systems.**

Space Climate or Climate in near-Earth space is characterized for long-term observations of space environment.

Space Weather refers for short-term , very dynamic and highly variable conditions in the geo-space environment.

- Impact on Solar Physics (with consequences even for fundamental particle physics; for example neutrino oscillation problem: SNO, Canada)
- It is important in order to improve our understanding of the Earth's climate and weather in relation with (some time) controversial problem of signatures of solar activity variability in meteorological parameters, Earth's atmosphere chemistry and long term trends in Earth's Climate.
- Big impact on space technology. Need to have Space Weather now cast and forecast . The very complex radiation effects on spacecraft systems and instruments, end even on Earth technology are influenced by Space Weather induced variations in the Earth's space environment

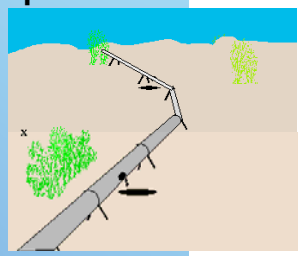
Disrupted Systems



Satellites

Increased drag on satellites in space, causing them to slow and change orbit

energetic solar particles. Energetic solar particles can cause physical damage to microchips and can change software commands. Another problem for satellite operators is differential charging



Pipelines

Geologic Exploration

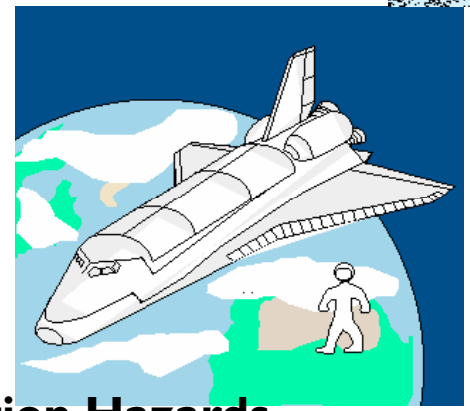
Climate **Biology**

Communications

Many communication systems utilize the ionosphere to reflect radio signals over long distances. Ionospheric storms can affect radio communication at all latitudes.

Electric Power

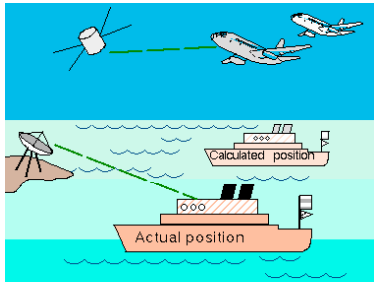
On March 13, 1989, in Montreal, Quebec, 6 million people were without commercial electric power for 9 hours as a result of a huge geomagnetic storm.



Radiation Hazards

Intense solar flares release very-high-energy particles that can be injurious to humans

Solar proton events can also produce elevated radiation aboard aircraft flying at high altitudes



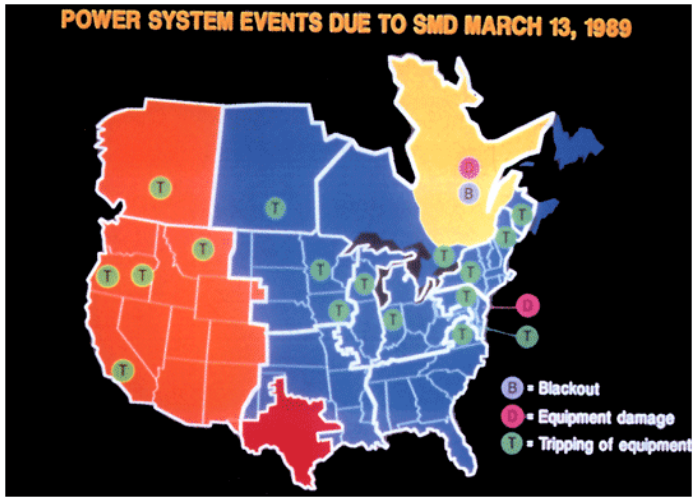
Navigation Systems

GROUND EFFECTS OF SPACE WEATHER

Large electric currents are continuously flowing in the magnetosphere and ionosphere. When hitting the magnetosphere, a disturbance in the solar wind produces a change in the current system, in which the magnetospheric-ionospheric coupling plays an important role. The geomagnetic field brings the disturbance in particular to high latitudes resulting in visible auroras and in an intense ionospheric current system.

The variations of magnetospheric and ionospheric currents are seen as geomagnetic disturbances or storms at the Earth's surface. In accordance with the basic electromagnetic theory, a geomagnetic variation is accompanied by a geoelectric field. The structure and intensity of the geoelectric field is greatly dependent on the Earth's conductivity structure.

smaller conductivity implies larger electric fields



The problems caused to power grids are due to a half-cycle saturation of transformers resulting from GIC. This means that a transformer which normally operates with a very small exciting current starts to draw an even hundred times larger current which results in a large asymmetry, and the transformer operates beyond the design limits.

Railway signalling systems in Swede were affected by a geomagnetic storm in July 1982. Figure credits: Andrew Pam.

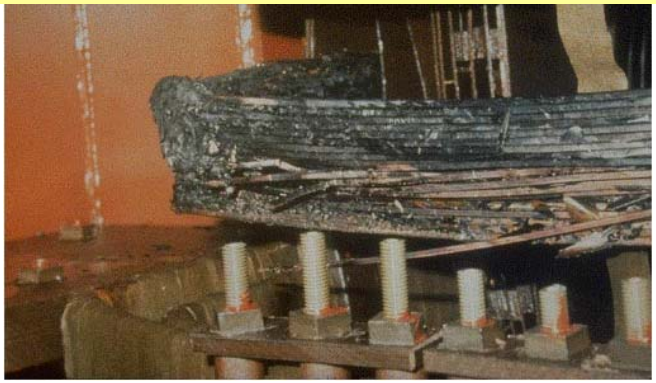
High-voltage power transmission systems are affected by geomagnetic disturbances.

First effects of GIC were experienced on telegraph equipment.

transformers located at corners of a power system suffer from large GIC values. Also, long transmission lines carry larger GIC. The problems caused to power grids are due to a half-cycle saturation of transformers resulting from GIC. This means that a transformer which normally operates with a very small exciting current starts to draw an even hundred times larger current which results in a large asymmetry, and the transformer operates beyond the design limits.

Image about the installation of a sea cable.

Magnetic surveys used in oil and gas exploration are disturbed by geomagnetic variations.



Damaged transformer windings.

Problem definition

The main purpose of this work is to extract interrelations between different variables related to geomagnetic phenomena.

Also, the geomagnetic activity is predicted.

1) we choose D_{st} index

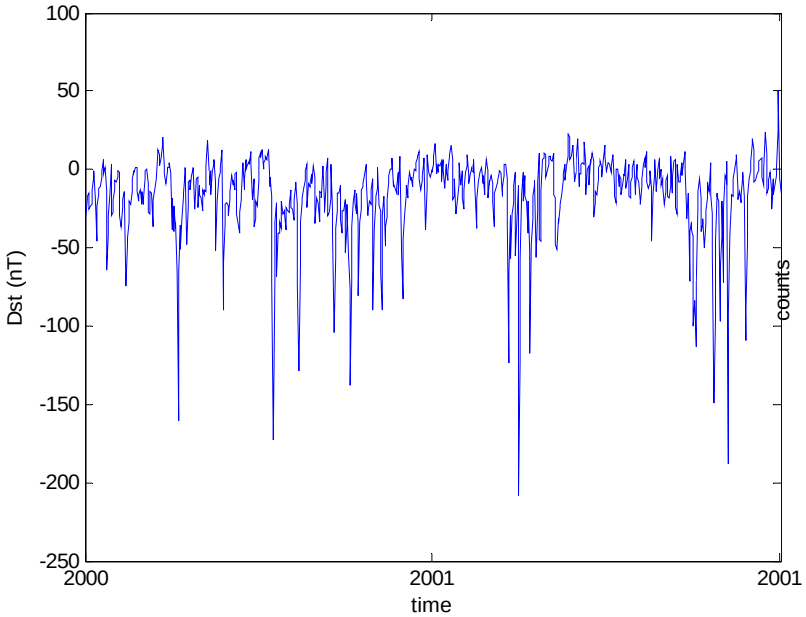
2) we calculate correlation coefficient and Mutual Information between D_{st} and its past measurements.

3) we calculate these measures between D_{st} and Cosmic Ray Intensity (CRI).

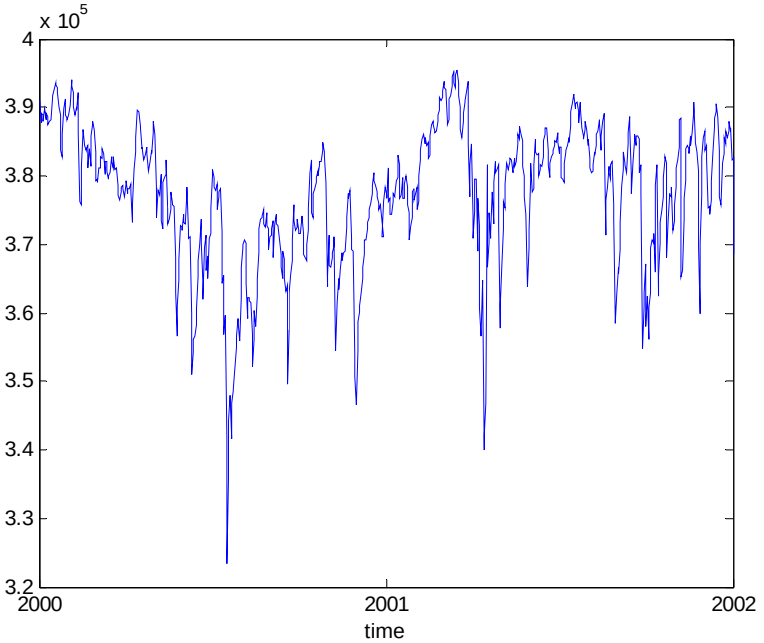
The utilized CRI data is from Terre Adelie station. Our study encompasses the collected time series between Jan. 1, 2000 and Dec. 31, 2001.

Simulation Results

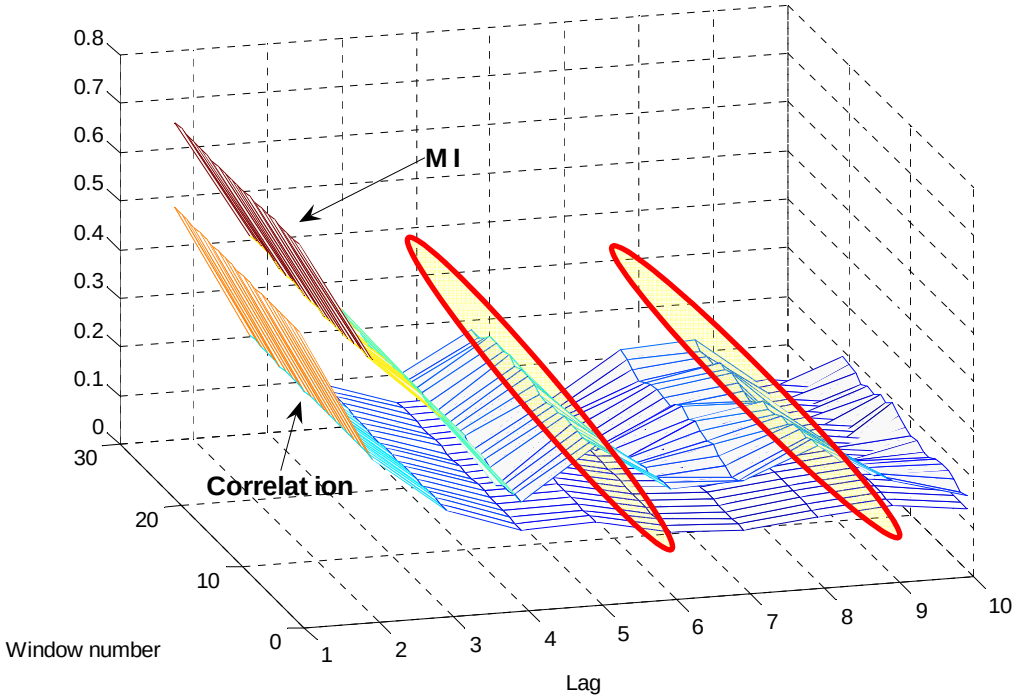
Daily measurements of D_{st} for 2000-2001



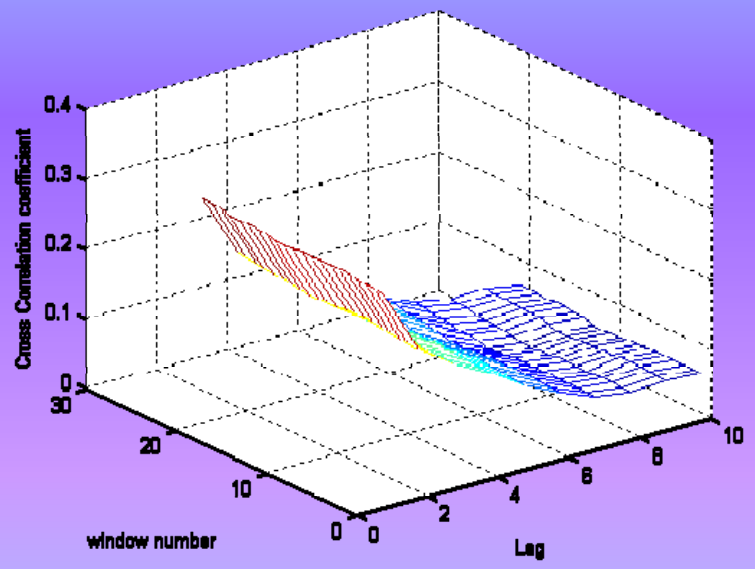
Daily measurements of CRI for 2000-2001



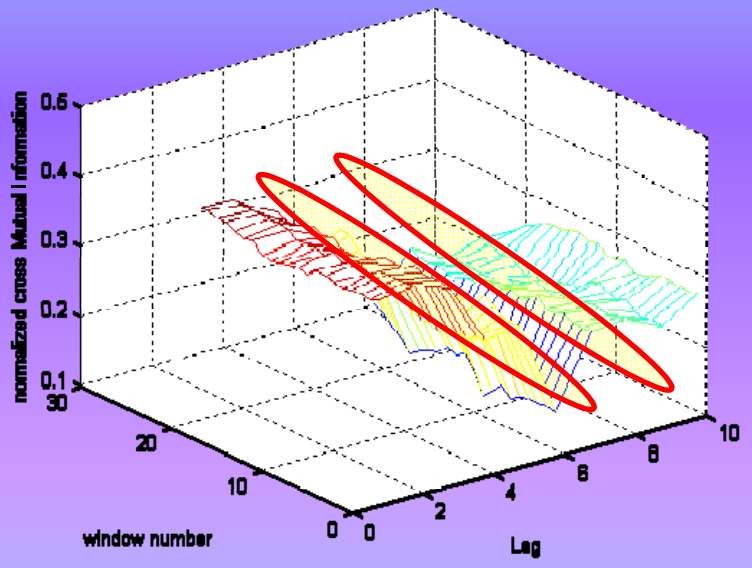
Auto Correlation coefficient values and Normalized Mutual Information between D_{st} and its past measurements



Cross Correlation coefficient values between D_{st} and CRI



Normalized cross Mutual Information between D_{st} and CRI

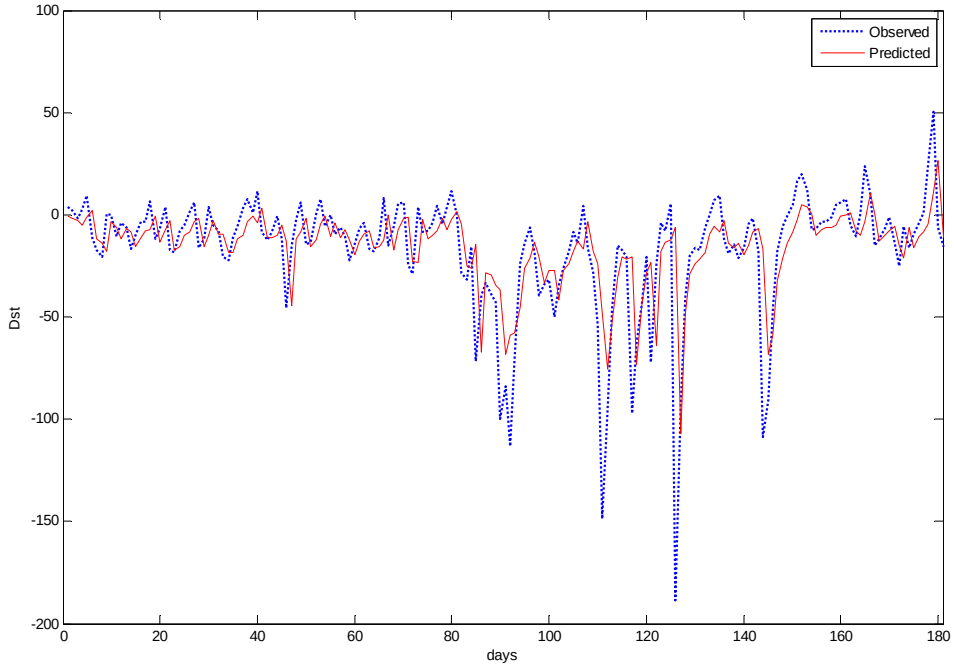


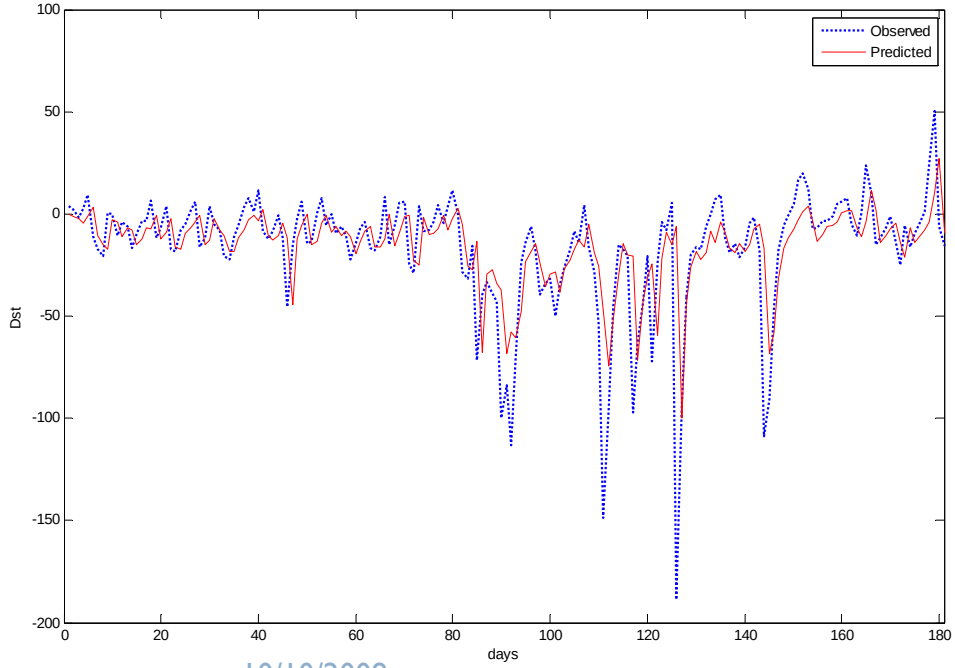
Also, for prediction stage we use Locally Linear Neurofuzzy model

Prediction Results:

Corr→

MI →





Compare the results

Input selection method (attention)	Prediction method	MSE
Corr.	3 neuron LLNF	0.6215
MI	3 neuron LLNF	0.6197

Conclusion

- The necessity for use of a focusing method like attention control is highlited.
- A framework for attention control based on information theoretical indices is proposed.
- The two measures of MI and Corr. is compared from stand point of computational costs and the value of information point of view.
- By employing information theoretical attention control, small number of inputs are selected for modeling or prediction of D_{st}.

- With respect to the computational costs, and constrains which is derived from limited computational power, and desired accuracy a subset of inputs and features using correlation or MI can be selected.

- Thanks for your attention,

Any Question ???