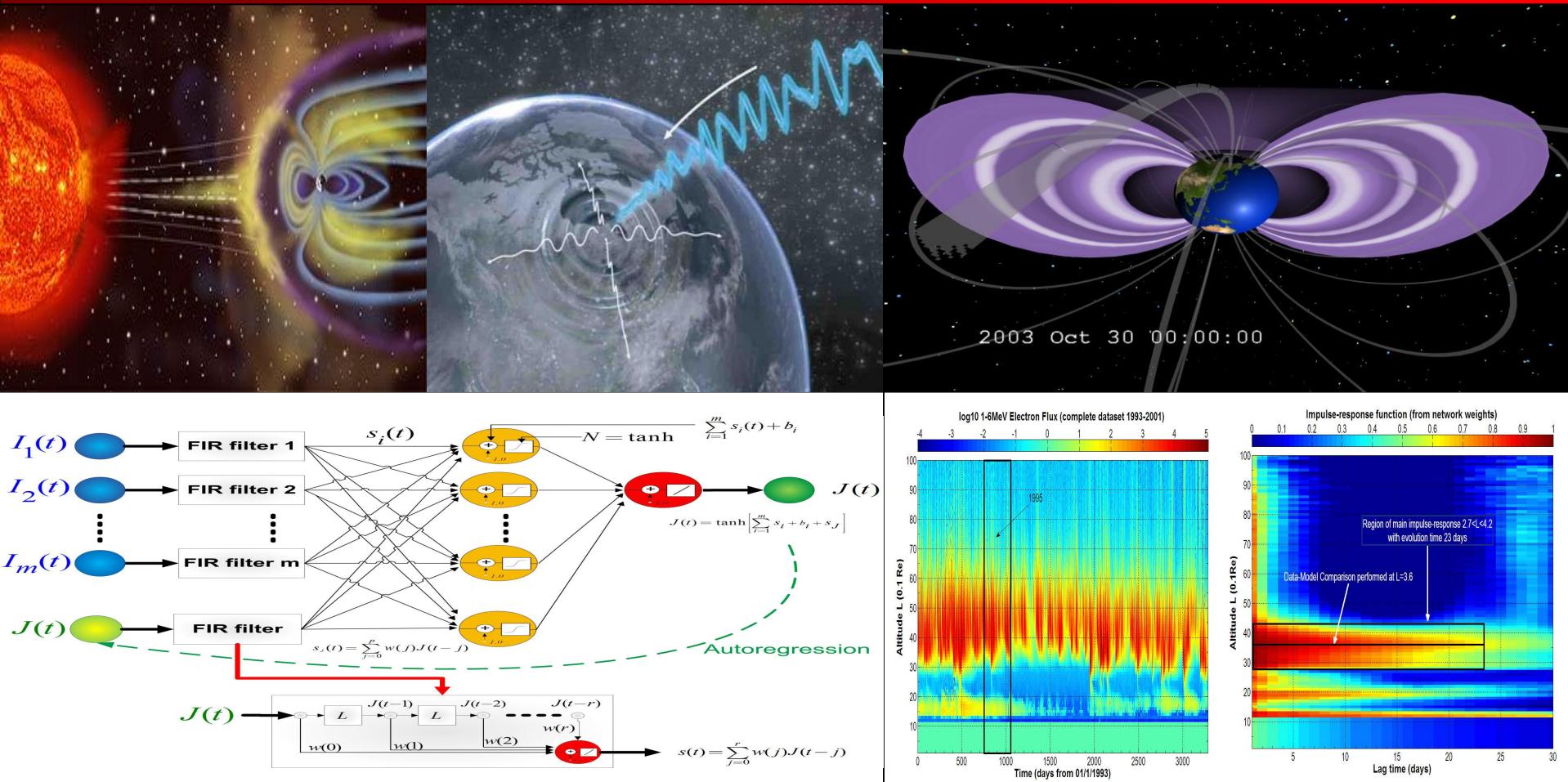


# How well can machines be trained to predict space weather storms?



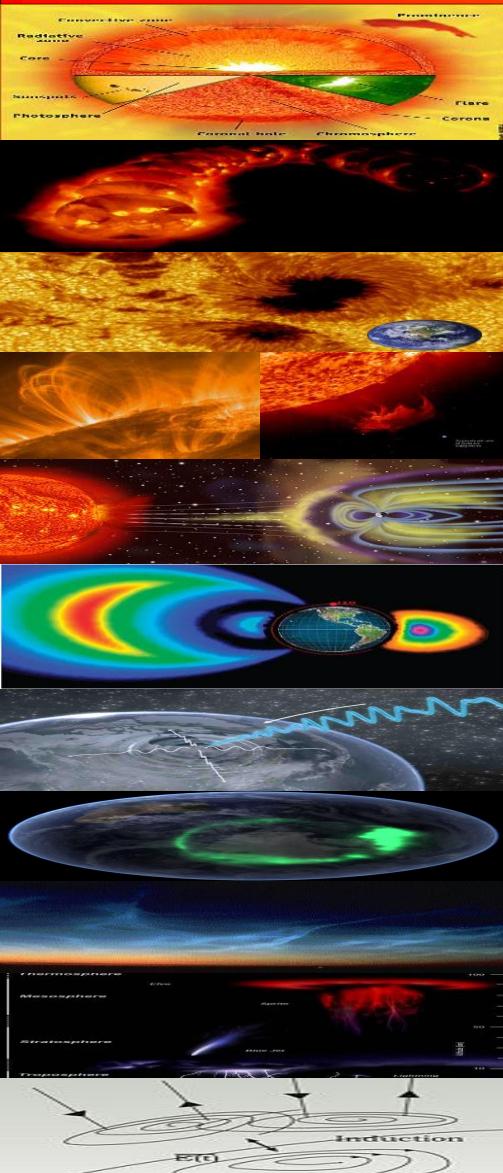
Michael Taylor

[michael@space.noa.gr](mailto:michael@space.noa.gr)

<http://patternizer.wordpress.com>



# Space Weather Storms: chain of events



**The Solar Interior** (*600 million tons of matter turn into energy every second*)

**The Solar Activity Cycle** (11 years)

**Sunspots**

**Solar Flares and Coronal Mass Ejections (CMEs)**

**The Solar Wind** (*Billions of tons of high velocity ejected plasma*)

**Van Allen Radiation Belts** (*MeV energy particles trapped at 3-6Re*)

**Substorm Impact** (*Magnetic blast waves can find impact epicentres*)

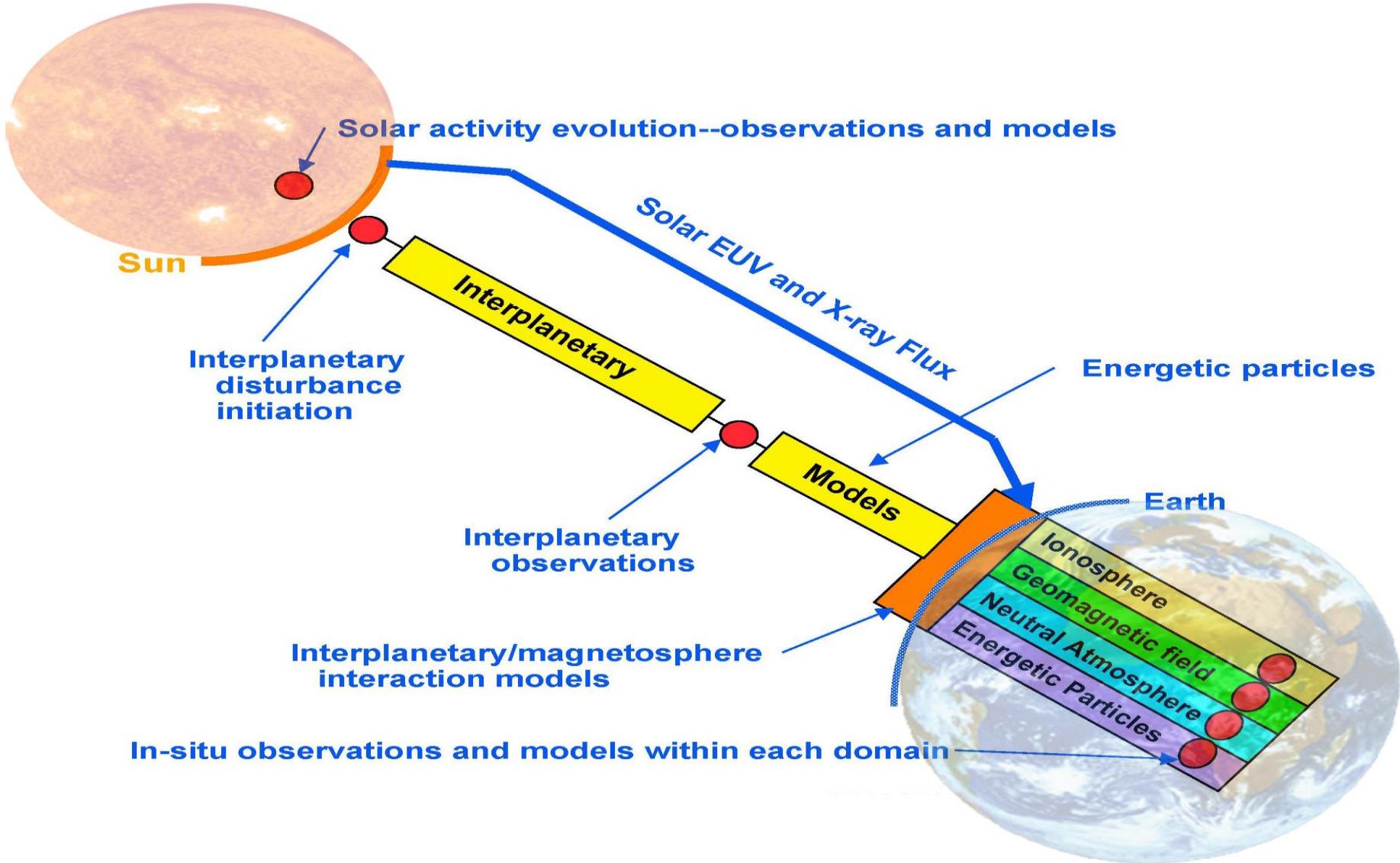
**The Aurorae**

**Plasma Clouds near the Earth** (*including upward "polar wind"*)

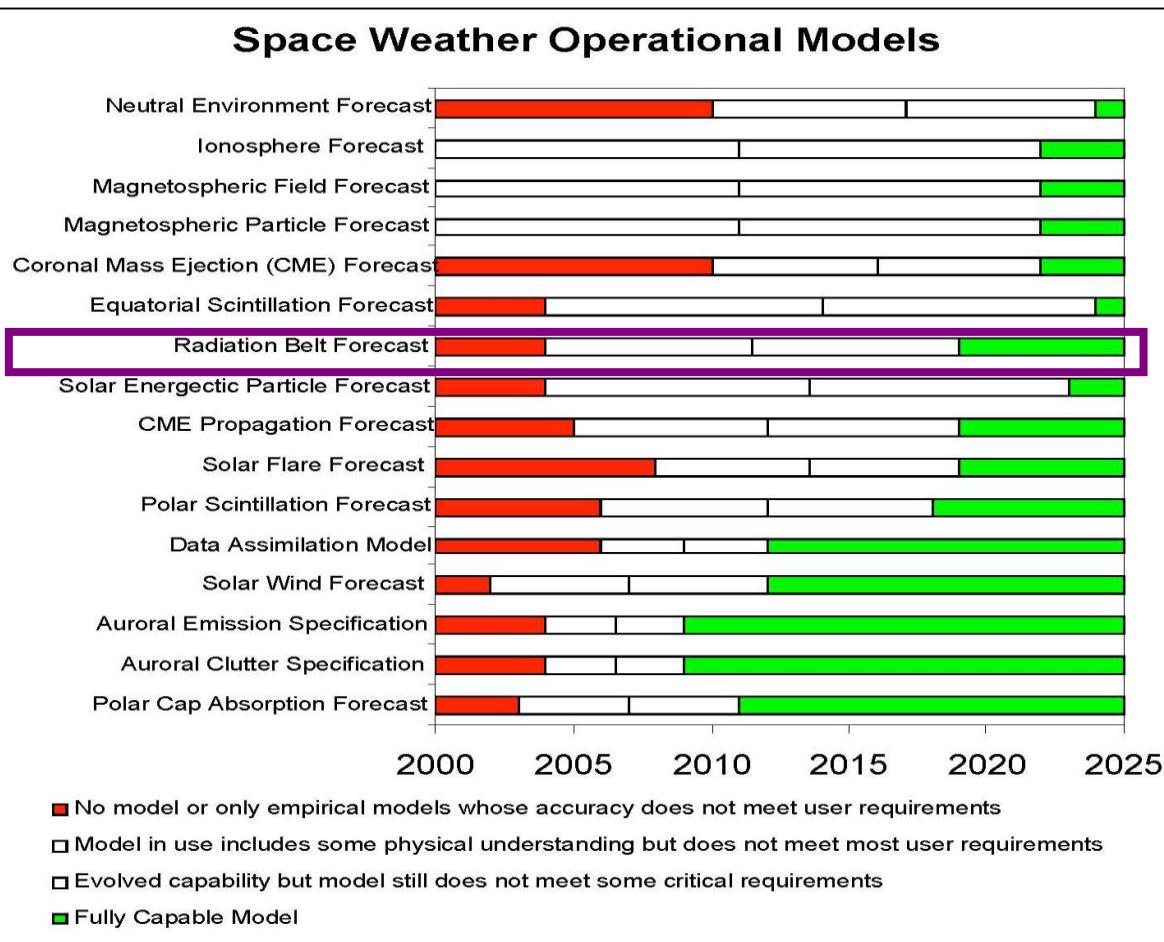
**Transient Luminous Events (TLEs) & Lightning** (*plasma-induced scintillations*)

**Geomagnetically Induced Currents (GICs)** (*due to downward-flowing plasma fields*)

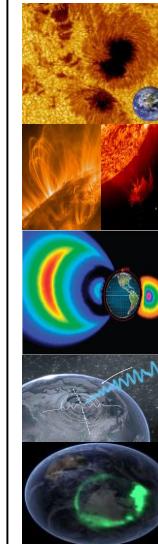
# Space Weather Modelling



# Forecast model status: 2000-2025



## New Data



Sunspots



CMEs



Radiation Belts



Substorm Impact

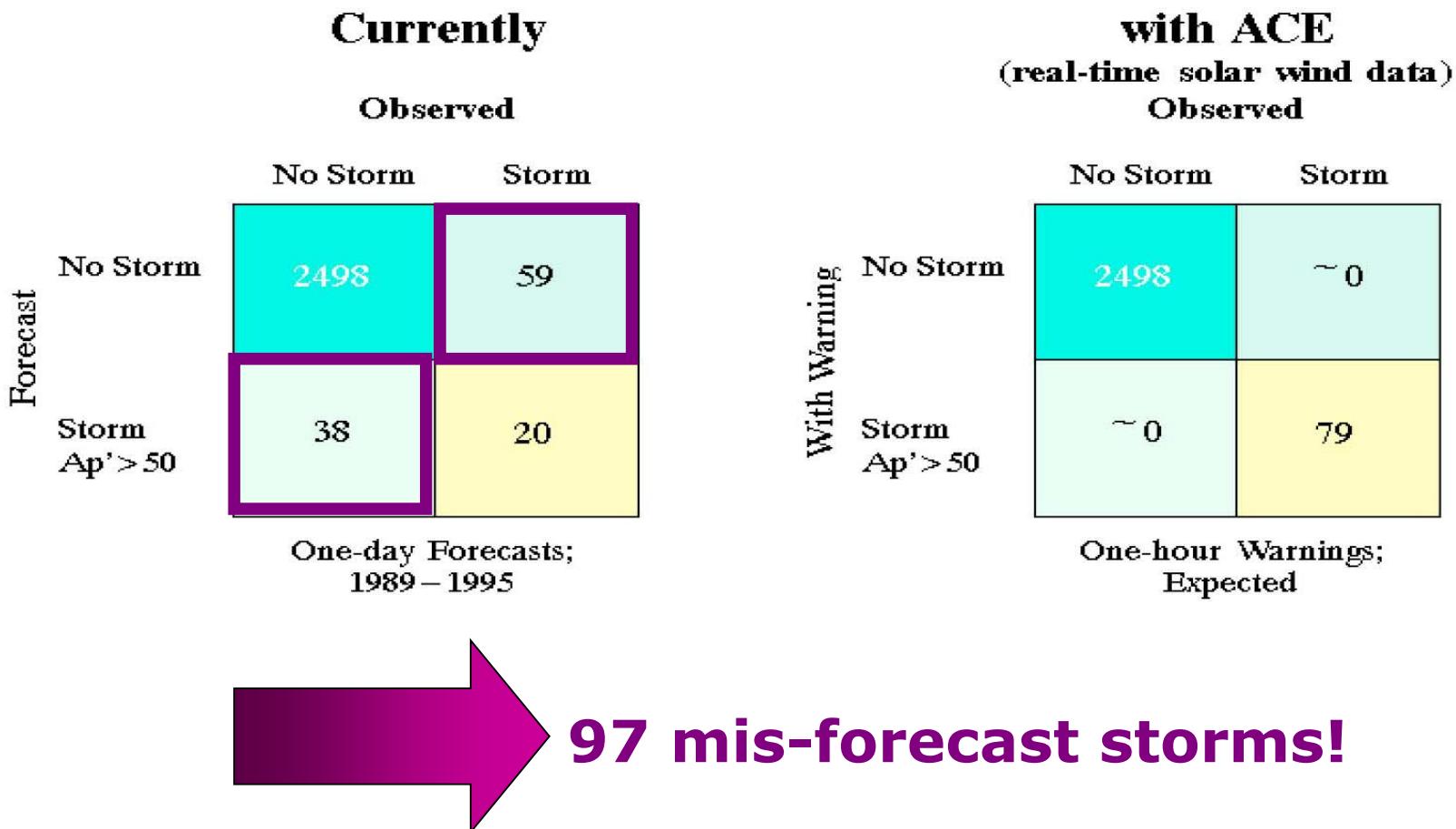


Aurorae



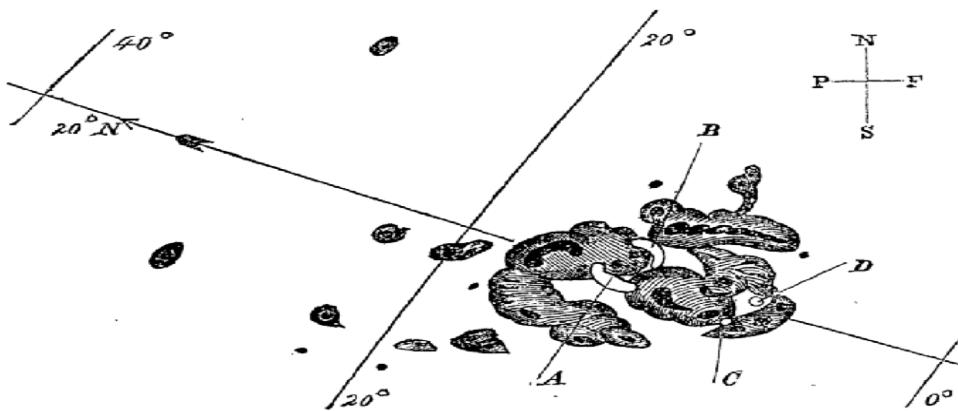
[Gary Heckman (2000) "Space Weather Forecasting Measurement and Modeling Requirements", Living with a Star Measurements Workshop, NASA/GSFC, February 9-10]

# 1-day forecasts (1989-1995)



[Gary Heckman (2000) "Space Weather Forecasting Measurement and Modeling Requirements", Living with a Star Measurements Workshop, NASA/GSFC, February 9-10]

# "The Superstorm" (August 28-2 September, 1859)



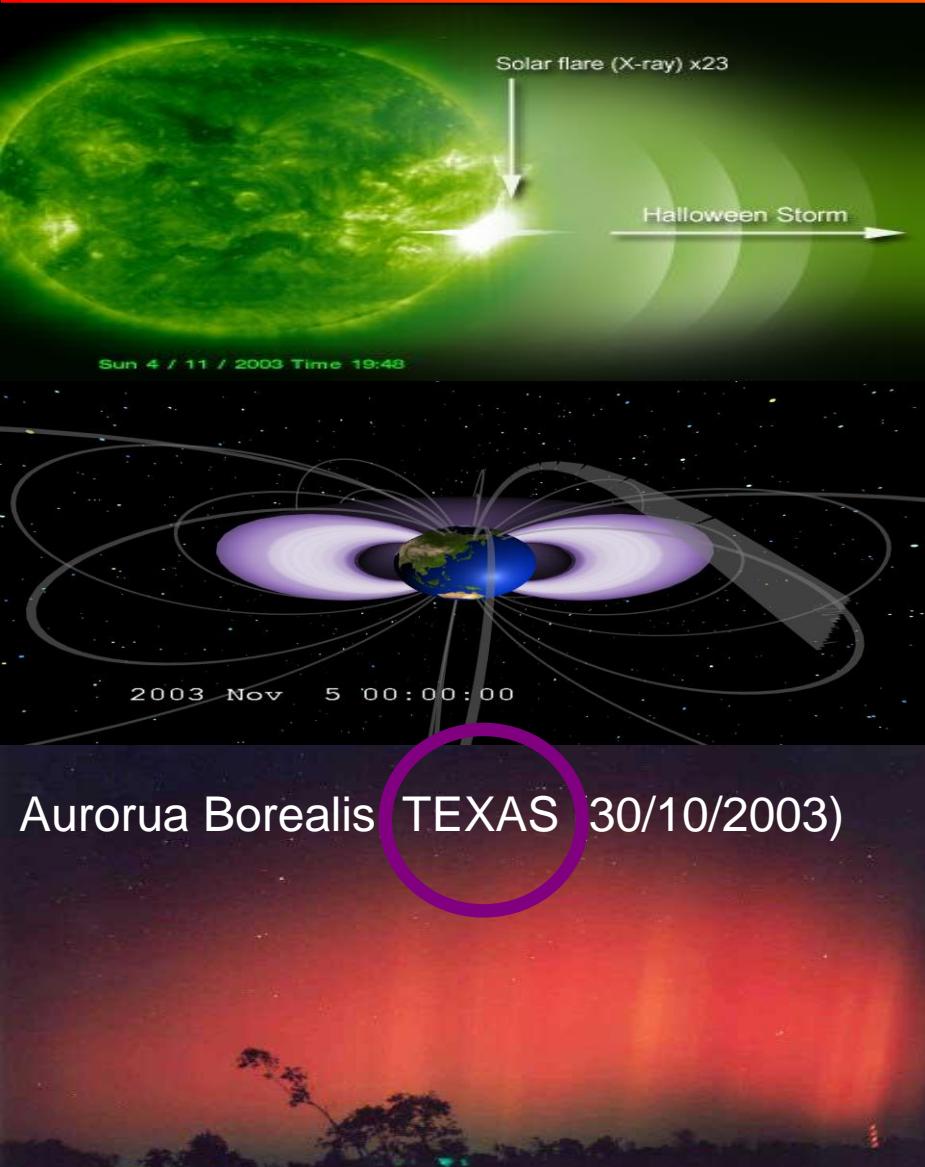
Carrington-Hodgson's naked-eye sketch of the **white light flare** that occurred on the sun on September 1<sup>st</sup> 1859 taking only 17 hours to hit Earth (more commonly 3-4 days)!

*"Towards half past eight o'clock a singular phenomenon took place. At first it was supposed that some great conflagration had taken place on the outskirts of the city, but it was soon recognized that no natural fires could produce this particular hue... Crowds of people gathered at the street corners, admiring and commenting upon the singular spectacle. Many took it to be the sign of some great disaster or important event, citing numerous instances when such warnings have been given. **Several old women were nearly frightened to death, thinking it announced the end of the world, and immediately took to saying their prayers.**"*

[ New Orleans Daily Picayune, p.5 ]

A recent report by the National Academy of Sciences found that if a similar storm occurred today, it could cause **\$1 to 2 trillion** in damages to society's high-tech infrastructure and require **four to ten years for complete recovery**. For comparison, Hurricane Katrina caused \$80 to 125 billion in damage.

# “The Halloween Storm” (29 October-4 November, 2003)



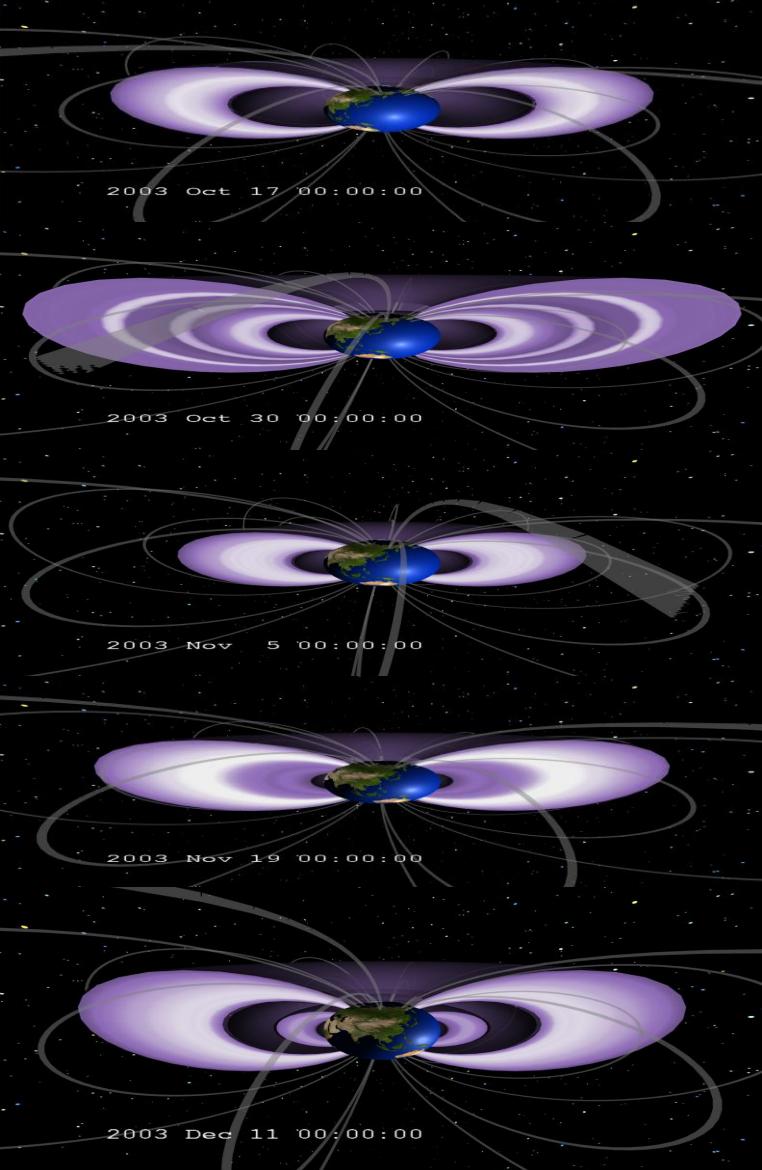
Astronauts hid deep within the body of the ISS, but reported radiation effects and ocular 'shooting stars'.

“The center of the outer Van Allen belt is usually about 12,000 miles to 16,000 miles away from Earth's surface. From 1-10 November, the outer belt was only about 6,000 miles from Earth.” Dan Baker, LASP

“We are able to understand and forecast more normal changes in the radiation belts using our present theoretical knowledge, but extreme events such as the Halloween storm are very hard to predict.” Xinlin Li, LASP

[ [ScienceDaily \(Dec. 9, 2003\)](#) ]

# “The Halloween Storm” (3D radiation belt evolution)



**60 days**

During the pre-storm time, the inner region of the belts has a relatively low particle flux and for this scaling of the data, a distinct inner belt is not visible as a separate structure.

On Halloween day, 3 radiation belts are detected.

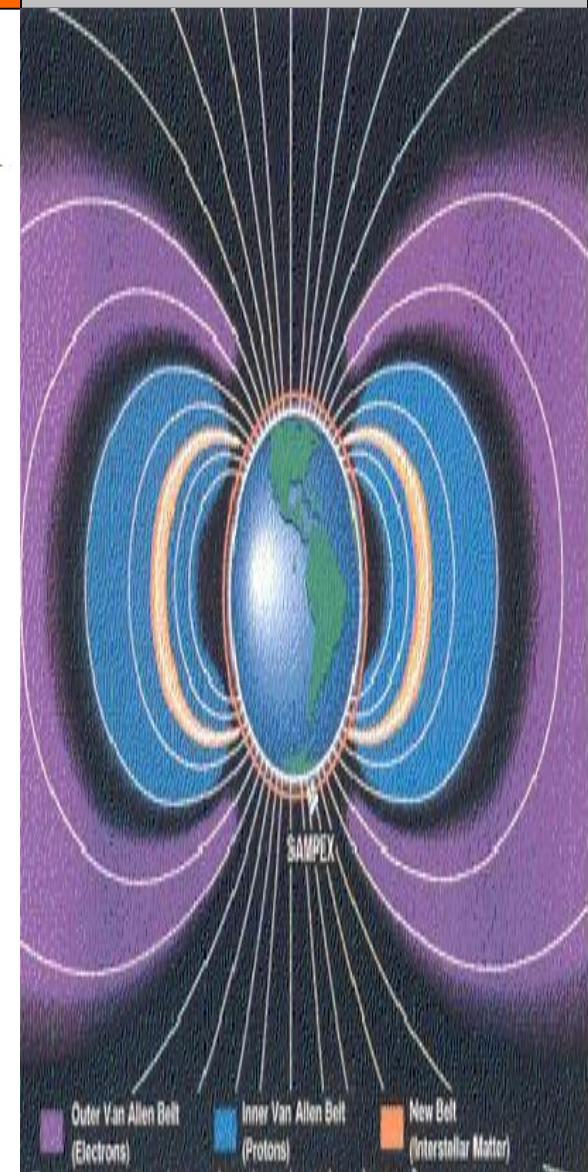
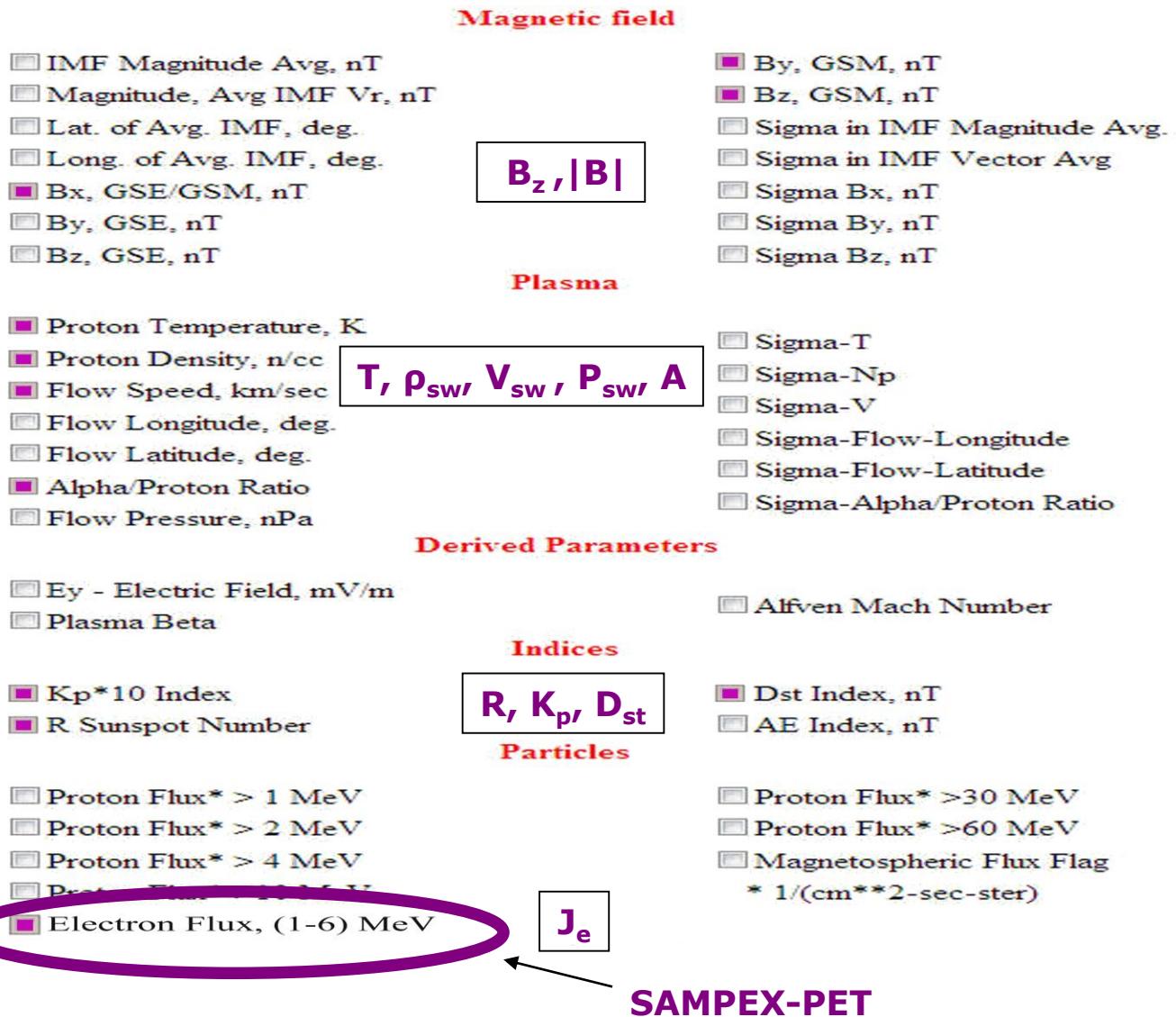
After the storm, the radiation belts are filled with particles and much closer to the Earth.

Over time, the belts relax, expanding back to their original locations but some residual particles remain close to the Earth.

Later, the outer belt has expanded even further and the inner belt retains some of its high particle flux.

[ NASA/GSFC Scientific Visualization Studio]

# 10 years of daily-averaged data for 11 SW,IP & MS parameters



How well can machines be trained to predict space weather storms?

Michael Taylor  
[michael@space.noa.gr](mailto:michael@space.noa.gr)

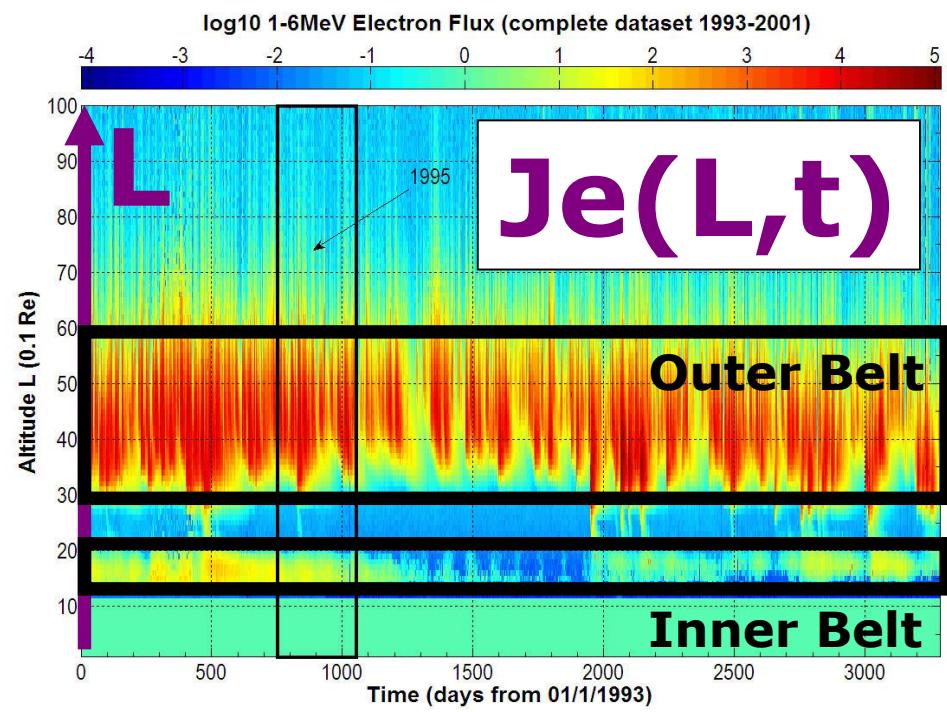
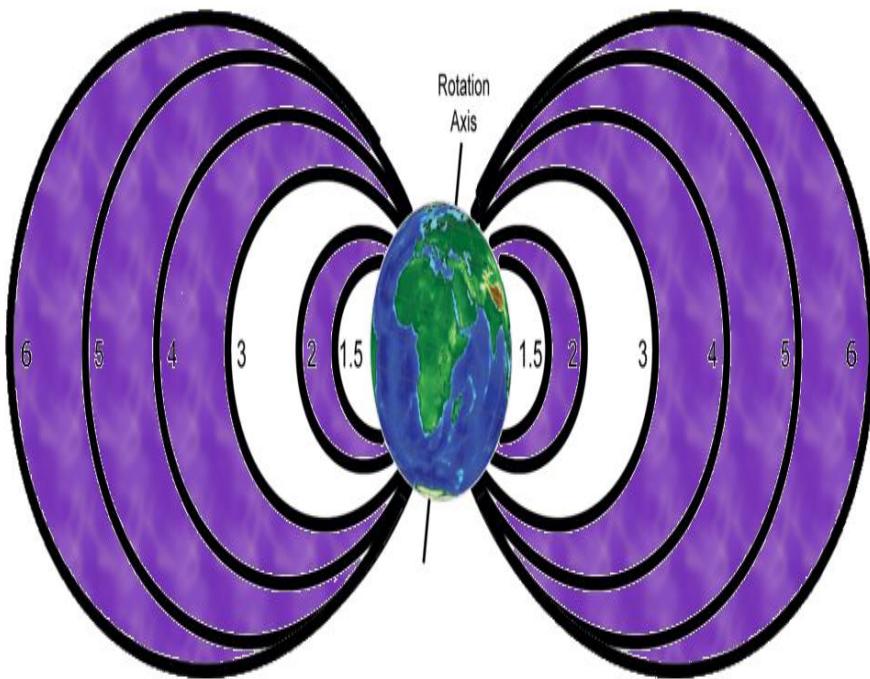
# Energetic particle flux $J$ in the radiation belts

The McIlwain “L-shell value” is a parameter describing a particular set of planetary magnetic field lines crossing the Earth's magnetic equator at a number of Earth-radii ( $R_E$ ) equal to the L

The Van Allen radiation belts roughly correspond to  $L=1.3 - 2$  (inner belt) and  $L=3 - 6$  (outer belt)

The plasmapause is typically around  $L=5$

Aurorae are most common around  $L=6$ , can reach  $L=4$  during moderate disturbances, and during the most severe geomagnetic storms, may approach  $L=2$



# In search of the physics...

$$J = f(V_{sw}, \rho_{sw}, P_{sw}, T, R, B_z, |B|, A, D_{st}, K_p)$$



**What is the “relevance list” of dominant variables ?**

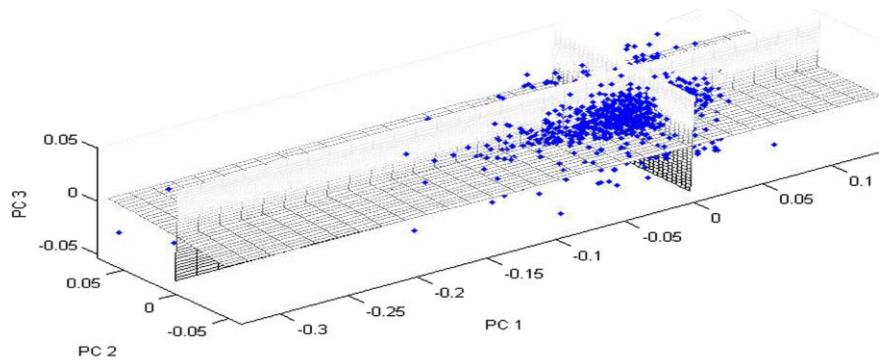


**What is the form of the function “f” ?**

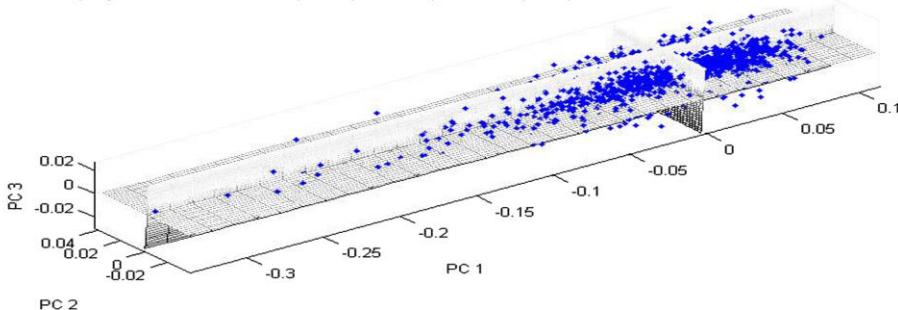
**How does “f” depend on the time-dependence of the relevant variables?**

# Strategy 1: Dimensionality Reduction (with a nonlinear PCA network)

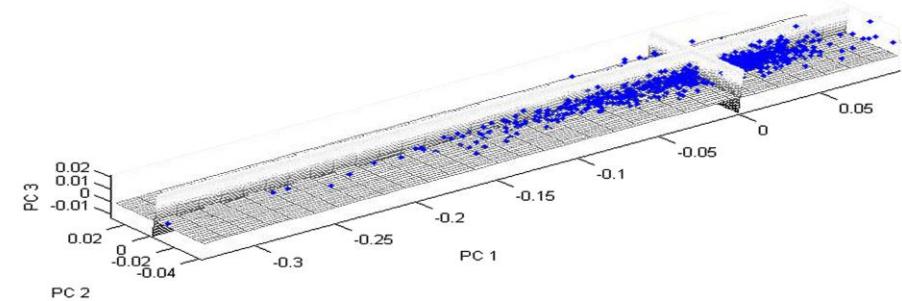
Data projection onto first 3 principal components (PCs) after 100 iterations



Data projection onto first 3 principal components (PCs) after 250 iterations

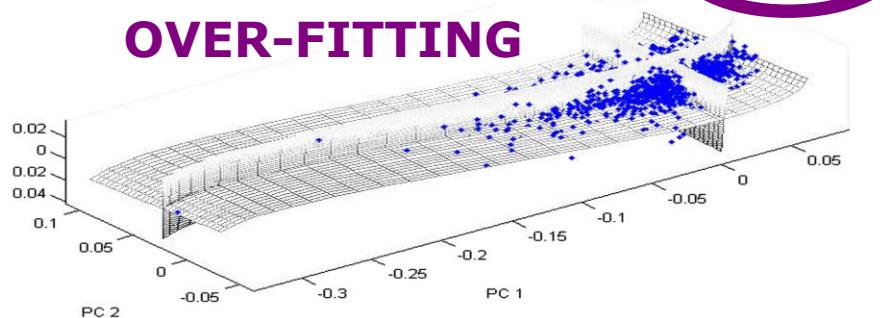


Data projection onto first 3 principal components (PCs) after 500 iterations



Data projection onto first 3 principal components (PCs) after 5000 iterations

**OVER-FITTING**



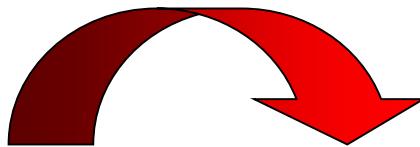
$$J = f(u_1, u_2, \dots, u_r) \quad r > 2$$

**Permutations**  $> p(12,3) = \frac{12!}{(12-3)!} = 1320$

**Conclusion:** 3 components are not enough – no chance of finding the dominant variables by trial and error → **Strategy 2!**

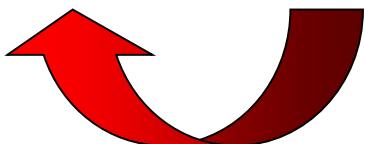
## Strategy 2: Dimensional Analysis

|                | $M$         | $L$ | $T$ | $Q$ | $\theta$ |
|----------------|-------------|-----|-----|-----|----------|
| $\mathbf{D} =$ | $J$         | 1   | 0   | -3  | 0        |
|                | $B_z$       | 1   | 0   | -1  | -1       |
|                | $D_{st}$    | 1   | 0   | -1  | -1       |
|                | $ B $       | 1   | 0   | -1  | -1       |
|                | $V_{sw}$    | 0   | 1   | -1  | 0        |
|                | $\rho_{sw}$ | 1   | -3  | 0   | 0        |
|                | $\mu_0$     | 1   | 1   | 0   | -2       |
|                | $k_B$       | 1   | 2   | -2  | 0        |
|                | $T$         | 0   | 0   | 0   | 1        |



$$J = V_{sw} P_{sw} f \left[ \frac{B_z}{C}, \frac{D_{st}}{C}, \frac{|B|}{C}, A, R, K_p \right]$$

where,  $C = \sqrt{\mu_0 P_{sw}}$



**(to be continued in Strategy 4)**

$$\mathbf{D}^\dagger = [ \mathbf{B}^\dagger \mid \mathbf{A}^\dagger ] = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -3 & 1 & 2 & 0 \\ -3 & -1 & -1 & -1 & -1 & 0 & 0 & -2 & 0 \\ 0 & -1 & -1 & -1 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

As a check on the completeness of the list of physical variables, we calculate the determinant  $|\mathbf{A}^\dagger| = -6 \neq 0$  as required for the invertibility of  $\mathbf{A}^\dagger$  [54]. The matrix product  $-(\mathbf{A}^\dagger)^{-1}\mathbf{B}^\dagger$  then gives the upper-diagonalised solution matrix  $\mathbf{E}$  [54],

$$\mathbf{E} = \begin{bmatrix} \mathbf{I} \\ -(\mathbf{A}^\dagger)^{-1}\mathbf{B}^\dagger \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ J & 1 & 0 & 0 \\ B_z & 0 & 1 & 0 \\ D_{st} & 0 & 0 & 1 \\ |B| & 0 & 0 & 1 \\ V_{sw} & -3 & -1 & -1 \\ \rho_{sw} & -1 & -1/2 & -1/2 & -1/2 \\ \mu_0 & 0 & -1/2 & -1/2 & -1/2 \\ k_B & 0 & 0 & 0 & 0 \\ T & 0 & 0 & 0 & 0 \end{bmatrix}$$

where  $\mathbf{I} = I_{mm}$  is the identity matrix. The  $m = 4$  different  $\pi$ -groups are then obtained from the similarity transform  $S$  [54]:

$$S : \pi_j = \prod_{i=1}^n x_i^{E_{ij}}$$

such that,

$$\begin{aligned} \pi_1 &= J^1 B_z^0 D_{st}^0 |B|^0 V_{sw}^{-3} \rho_{sw}^{-1} \mu_0^0 k_B^0 T^0 = \frac{J}{V_{sw} P_{sw}}, \\ \pi_2 &= J^0 B_z^1 D_{st}^0 |B|^0 V_{sw}^{-1} \rho_{sw}^{-1/2} \mu_0^{-1/2} k_B^0 T^0 = \frac{B_z}{\sqrt{\mu_0 P_{sw}}}, \\ \pi_3 &= J^0 B_z^0 D_{st}^1 |B|^0 V_{sw}^{-1} \rho_{sw}^{-1/2} \mu_0^{-1/2} k_B^0 T^0 = \frac{D_{st}}{\sqrt{\mu_0 P_{sw}}}, \\ \pi_4 &= J^0 B_z^0 D_{st}^0 |B|^1 V_{sw}^{-1} \rho_{sw}^{-1/2} \mu_0^{-1/2} k_B^0 T^0 = \frac{|B|}{\sqrt{\mu_0 P_{sw}}} \end{aligned}$$

Using the similarity transform method of:  
**[Taylor M et al (2008) Adv Stud. Theor. Phys, Vol 2(20):979 -995]**

## Strategy 3: Volterra Networks

**Volterra Theorem:** The Volterra series is the output of a nonlinear system and depends on the input to the system at all other times

[Volterra V (1887) in: Volterra V (1913) *Leçons sur les équations intégrales et les équations intégro-différentielles* Paris: Gauthier-Villars]



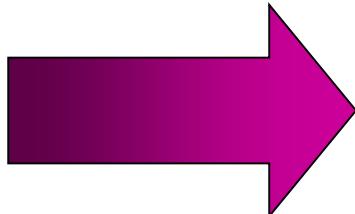
**Takens' Theorem:** Time-delay embedding allows the nonlinear dynamics to be reconstructed from a sequence of observations of the state of a dynamical system

[Takens (1981) *Dynamical Systems and Turbulence: Lecture Notes in Mathematics 898*:366-381]



**Hornik's Theorem:** A multilayer feed-forward network with a single hidden layer containing a finite number of hidden neurons and with an arbitrary activation function are universal approximators

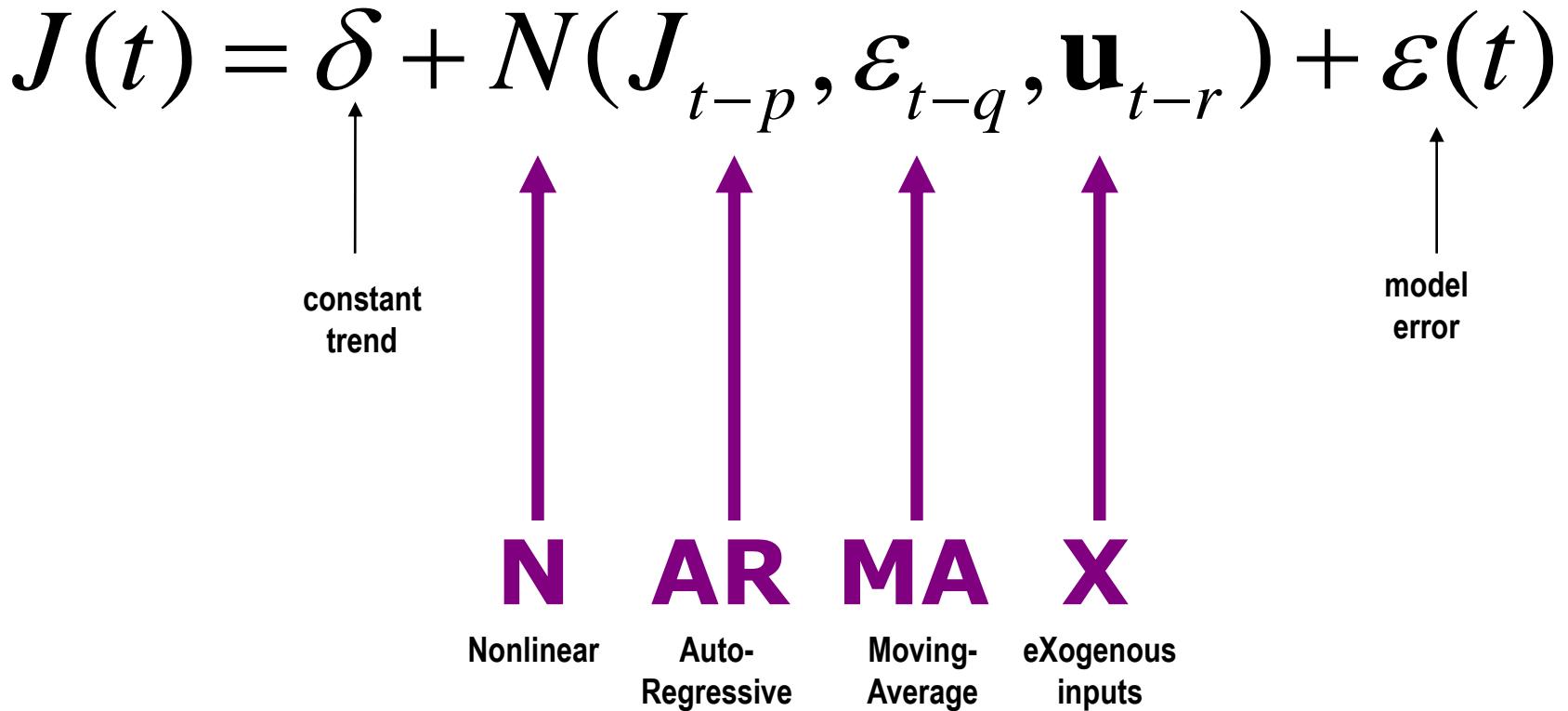
[Hornik (1991) *Approximation Capabilities of Multilayer Feedforward Networks*. *Neural Networks* 4]



## Nonlinear Volterra Neural Networks (= NARMAX networks)

[Taylor M, Daglis IA, Anastasiadis A, Vassiliadis D (2009) Proc 9th Hellenic Astrophysical Conference]

# NARMAX input-output models



# A taxonomy of NARMAX models

Table 1: A taxonomy of input-output models under the NARMAX class

| Function | Autoregression order | Moving-Average order | Inputs | Model                             |
|----------|----------------------|----------------------|--------|-----------------------------------|
| 1        | 1                    | 0                    | 0      | AR(1)=Random walk                 |
| 1        | $p$                  | 0                    | 0      | AR( $p$ )                         |
| 1        | $p$                  | 0                    | $r$    | ARX( $p, r$ )                     |
| 1        | 0                    | $q$                  | 0      | MA( $q$ )                         |
| 1        | 0                    | $\infty$             | 0      | MA( $\infty$ )=Wold Decomposition |
| 1        | 0                    | $q$                  | $r$    | MAX( $q, r$ )                     |
| 1        | $p$                  | $q$                  | 0      | ARMA( $p, q$ )                    |
| 1        | $p$                  | $q$                  | $r$    | ARMAX( $p, q, r$ )                |
| $F$      | $p$                  | 0                    | 0      | NAR( $p$ )                        |
| $F$      | $p$                  | 0                    | $r$    | NARX( $p, r$ )                    |
| $F$      | 0                    | $q$                  | 0      | NMA( $q$ )                        |
| $F$      | 0                    | $q$                  | $r$    | NMAX( $q, r$ )                    |
| $F$      | $p$                  | $q$                  | 0      | NARMA( $p, q$ )                   |
| $F$      | $p$                  | $q$                  | $r$    | NARMAX( $p, q, r$ )               |

**where  $F=N$  as determined by the architecture  
of the Volterra neural network**

[Taylor M et al (2010) A taxonomy of NARIMAX input-output models for nonlinear multiparameter modelling of the radiation belt energetic electron flux (*in preparation*)]

# The Modelling Approach

Space weather variables

$$J_{e,i}, V_{sw}, P_{sw}, \rho_{sw}, T, R, B_z, |B|, A, D_{st}, K_p$$

Observational time series data

$$J(t), V_{sw}(t), P_{sw}(t), \rho_{sw}(t), T(t), R(t), B_z(t), |B|(t), A(t), D_{st}(t), K_p(t)$$

Time series decomposition

$$J(t) = \delta + f[\Psi(t)] + \varepsilon(t)$$

Taxonomy of NARMAX( $p, q, r$ ) MIMO Models

1

$$\Psi_t = [\varphi_1 J_{t-1}, \dots, \varphi_p J_{t-p}; \theta_1 \varepsilon_{t-1}, \dots, \theta_q \varepsilon_{t-q}; \eta_1 u_{t-1}, \dots, \eta_r u_{t-r}]$$

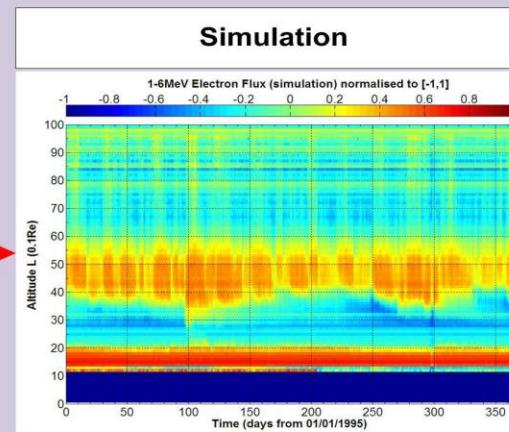
Nonlinear Models  
 $f=N=nonlinear$

2

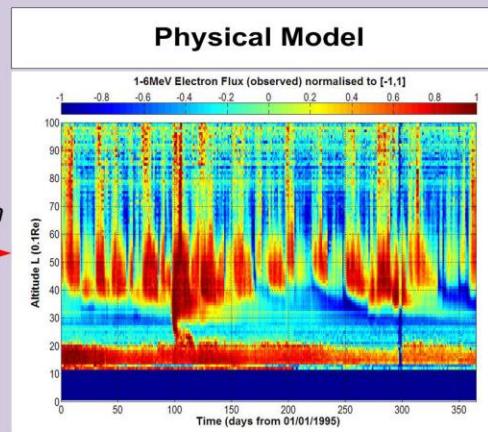
$$J(t) = \delta + N[\Psi(t)] + \varepsilon(t) \equiv \hat{J}(t) + \varepsilon(t)$$

Takens' Theorem  
Nonlinear Volterra Network  
(time-delay FIR network)

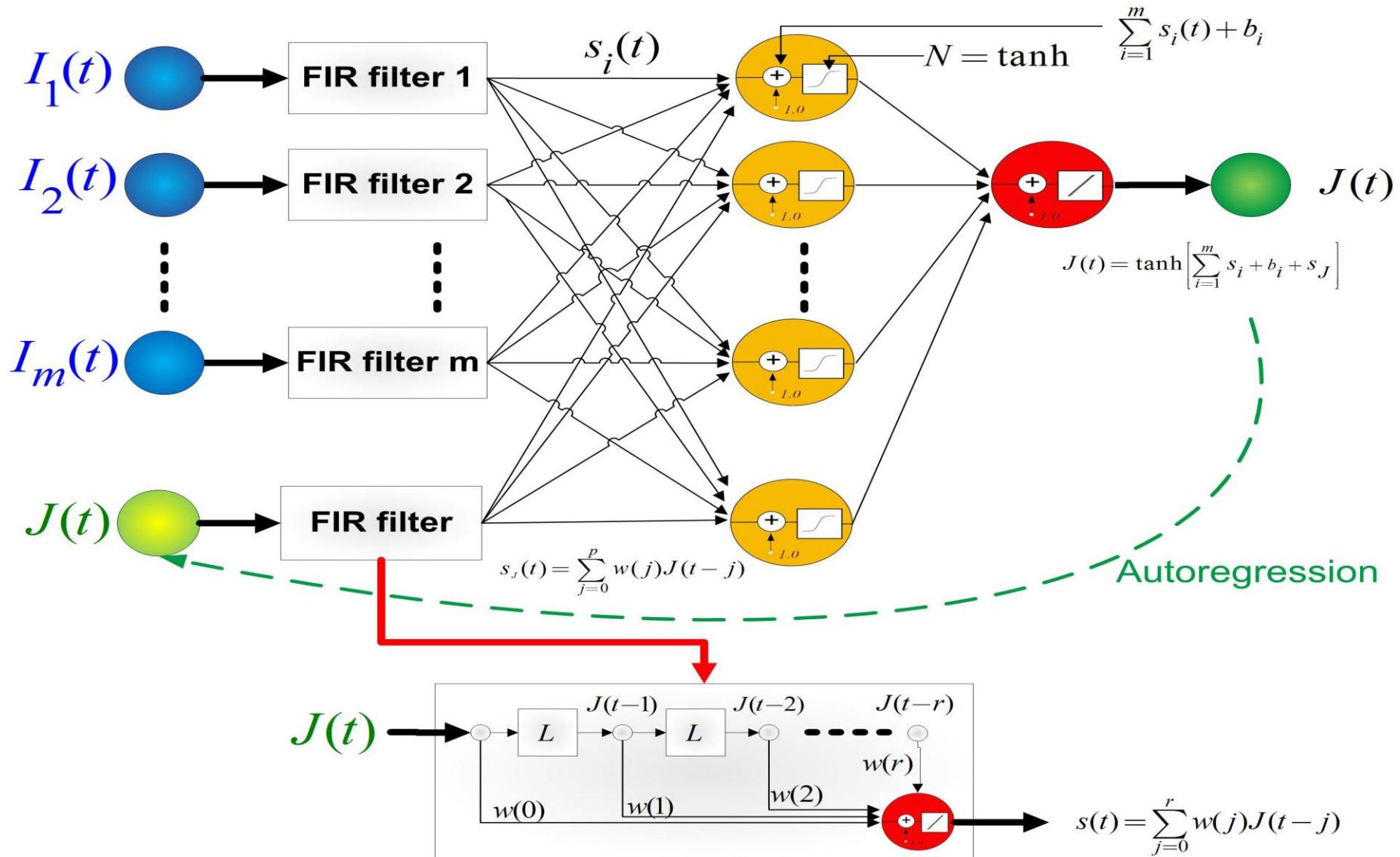
$$\hat{J}(t) = \delta + N[\Psi(t)]$$



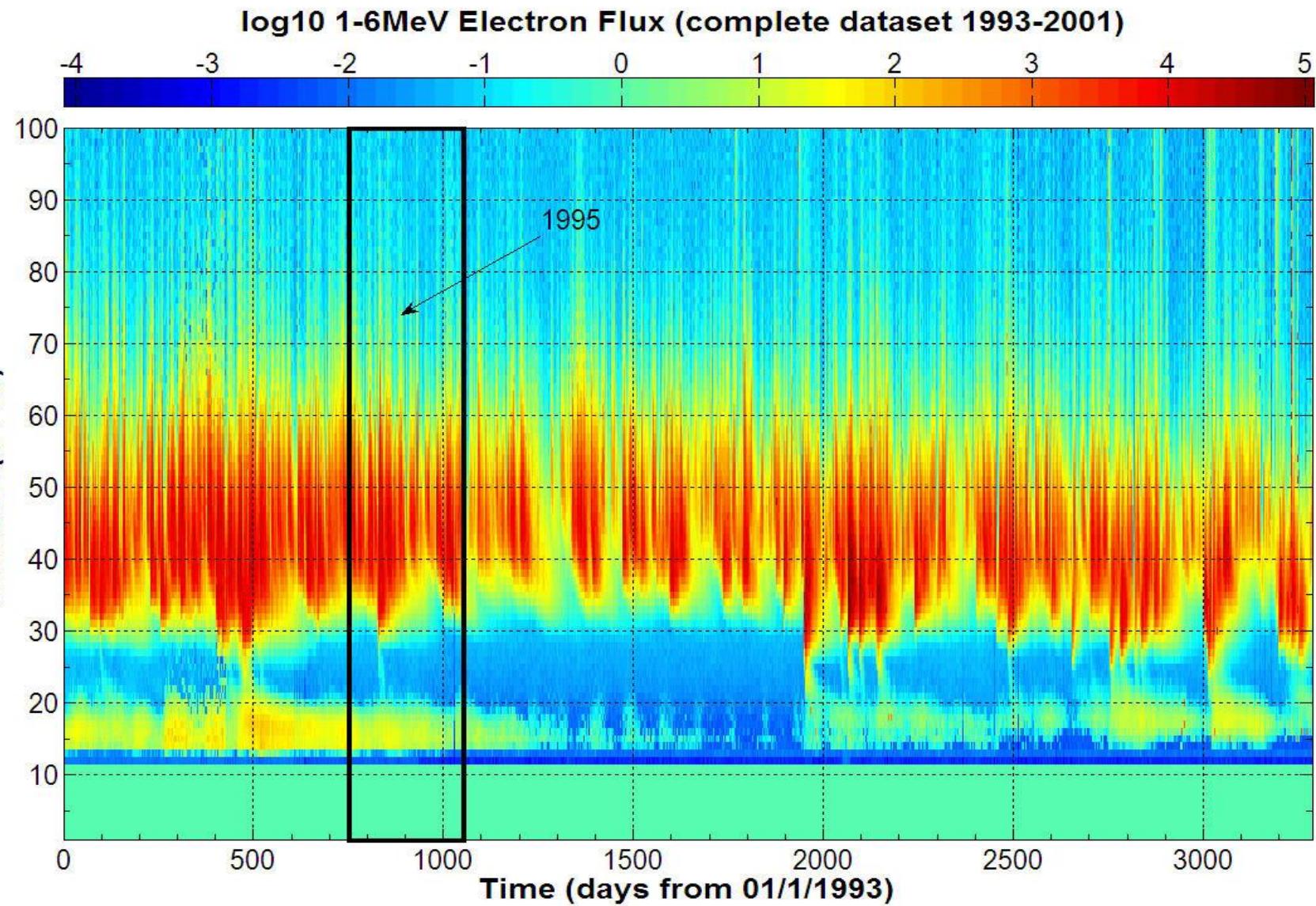
Comparison  
4  
 $J(t)$



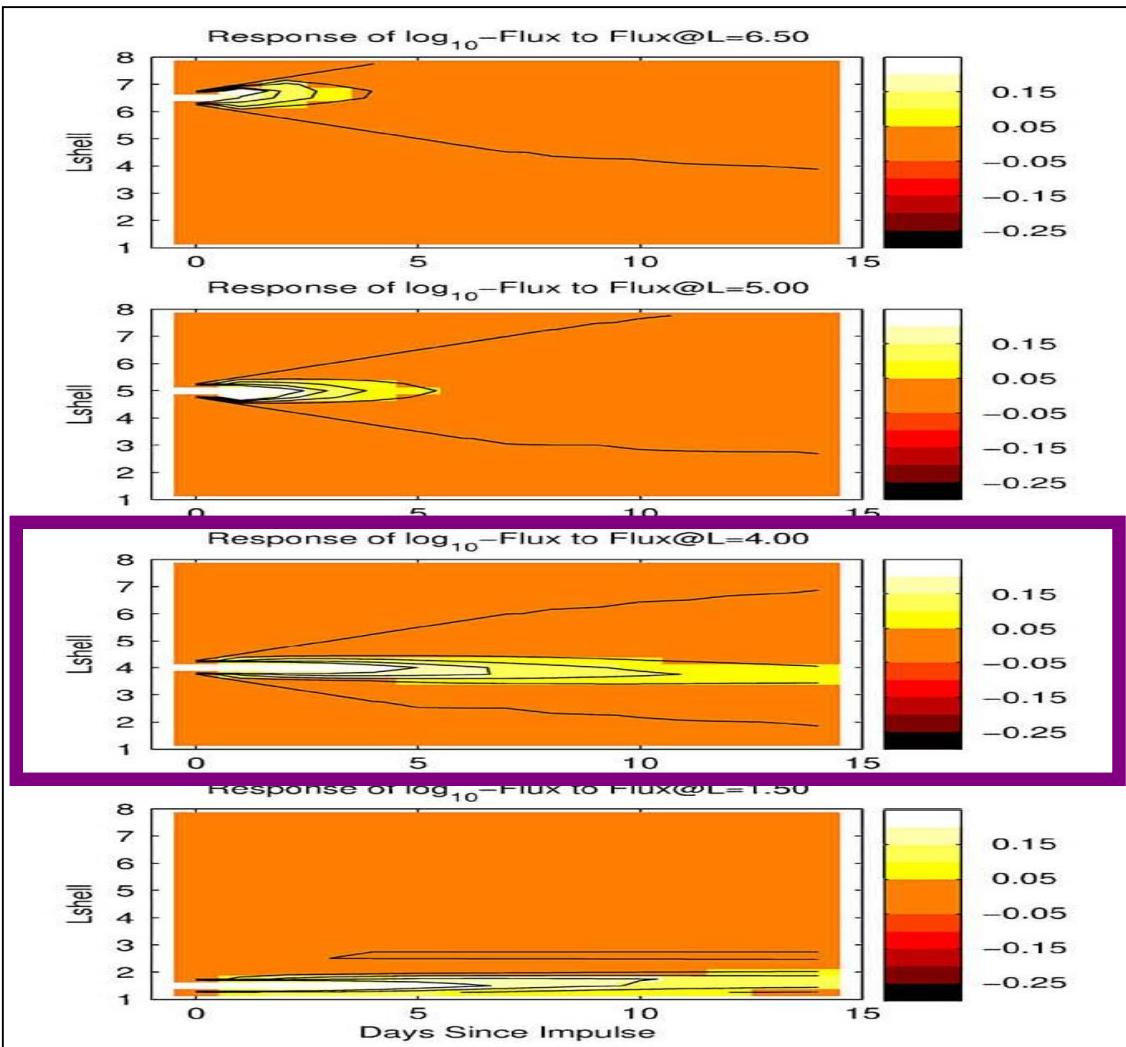
# Volterra (NARMAX) Neural Networks



# The Radiation Belt Electron Flux ( $L, t$ )



# State-space method



**State Vector**

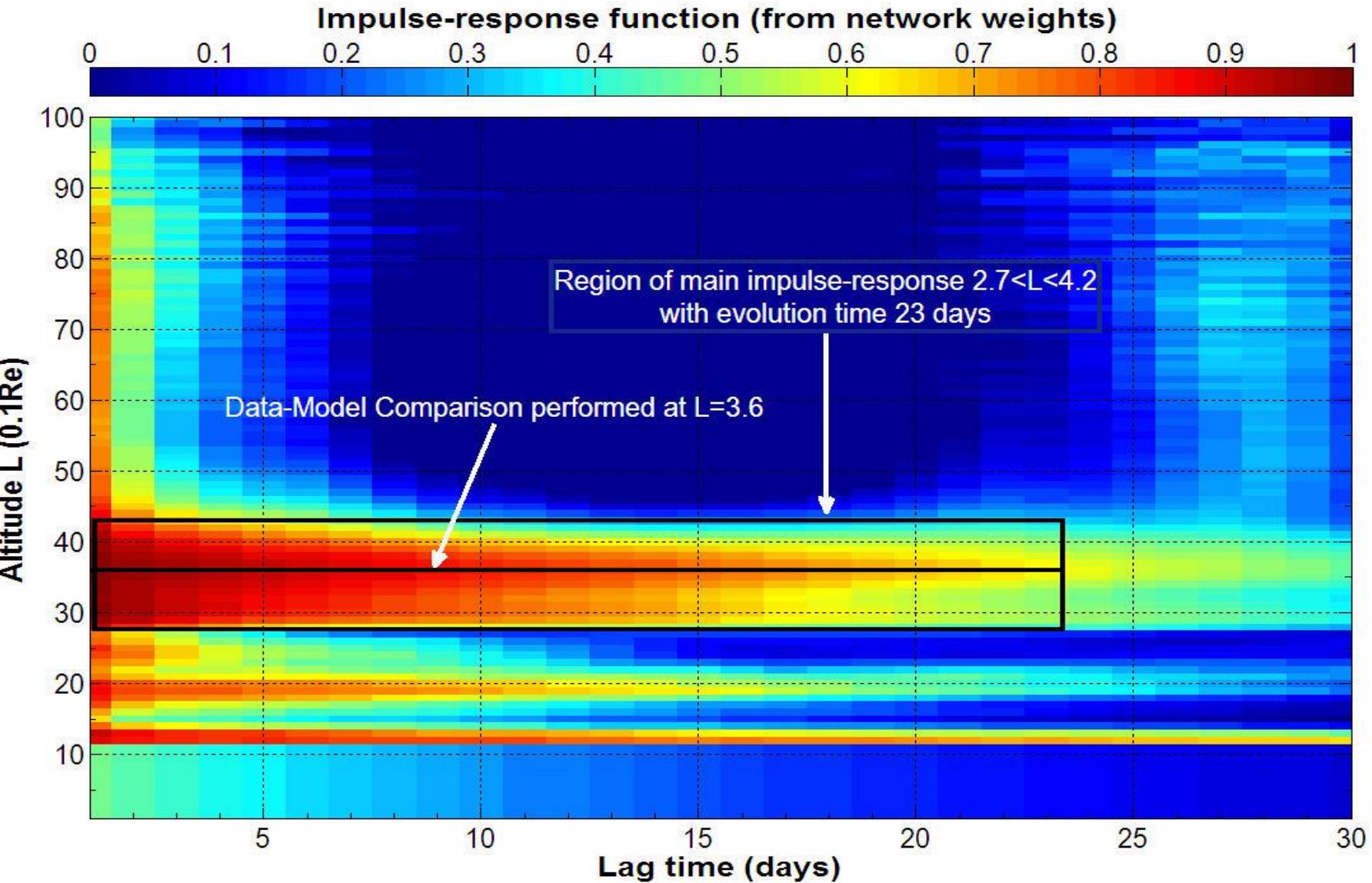
$$\mathbf{x}(t) = F_1 \left[ \mathbf{x}(t-1), e(t-1) \right]$$

$$\mathbf{y}(t) = F_2 \left[ \mathbf{x}(t) \right] + e(t)$$

**Output Vector**

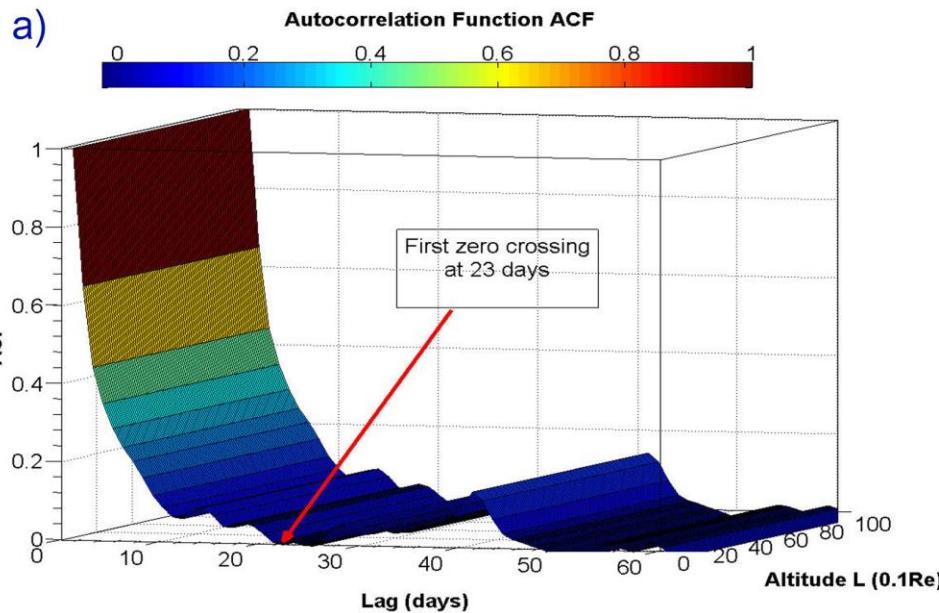
[Rigler EJ and Baker DN (2008) JASTP 70:1797–1809]

# NAR(23)-Model (1995) (Nonlinear Finite Impulse-Response)

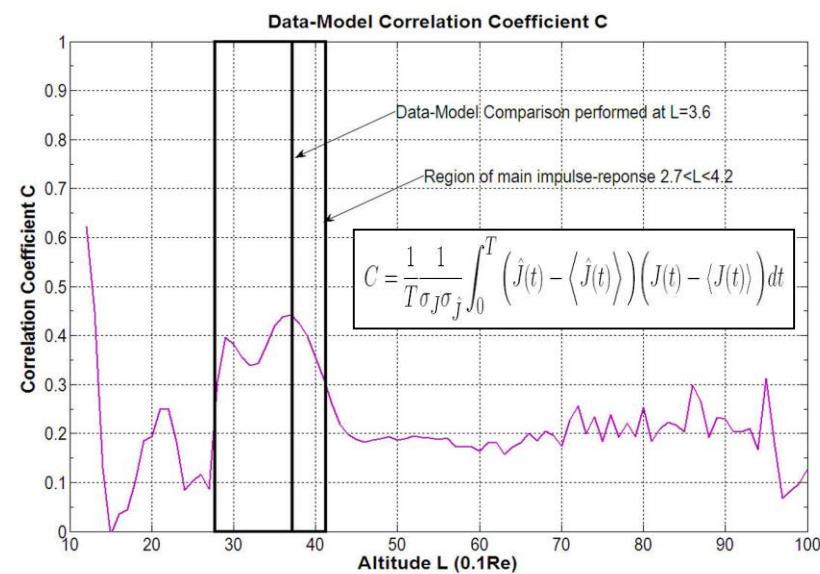


# NAR(23)-Model (1995) (Data-Model Correlation)

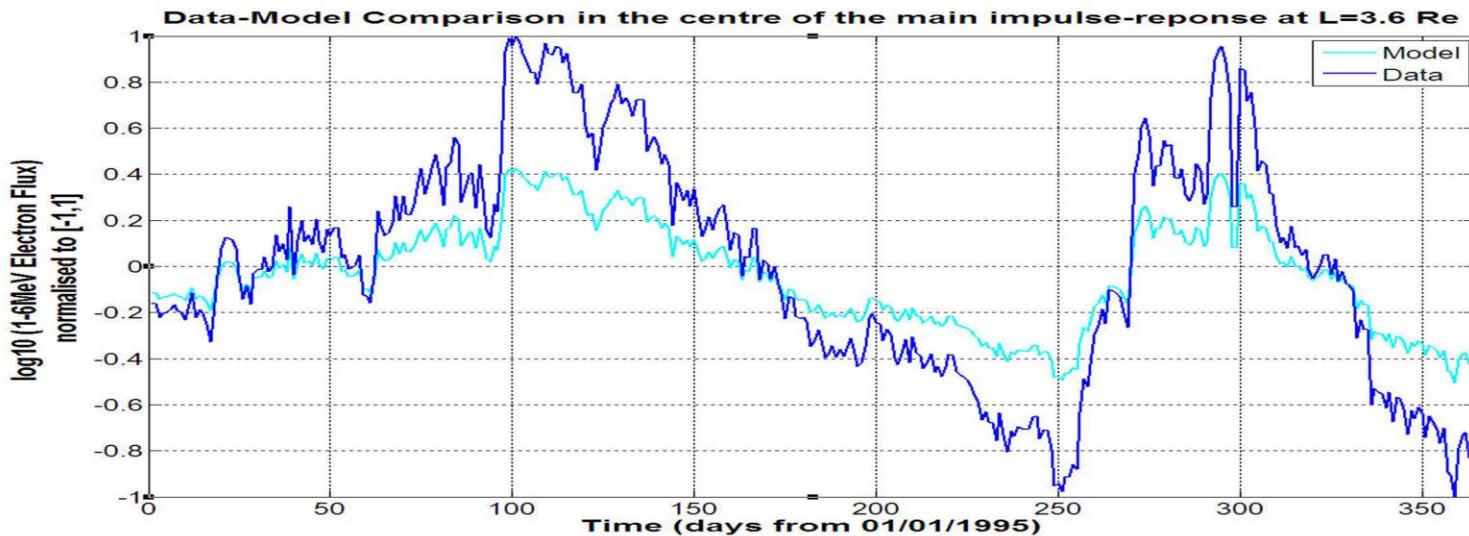
a)



b)

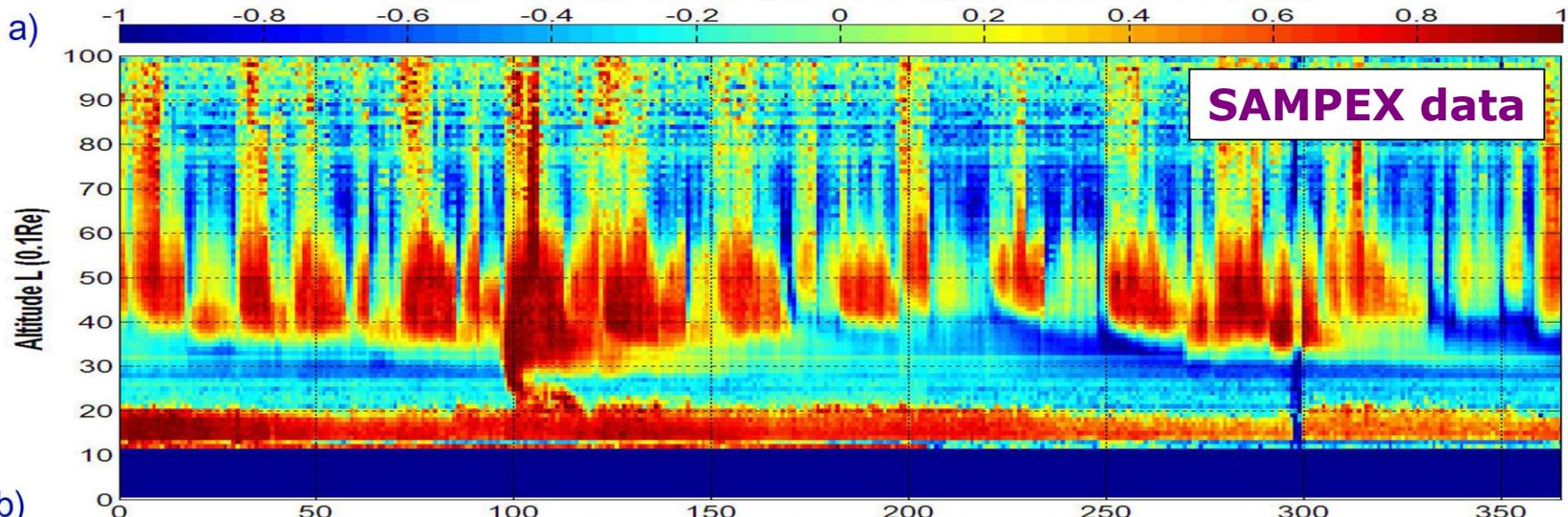


c)

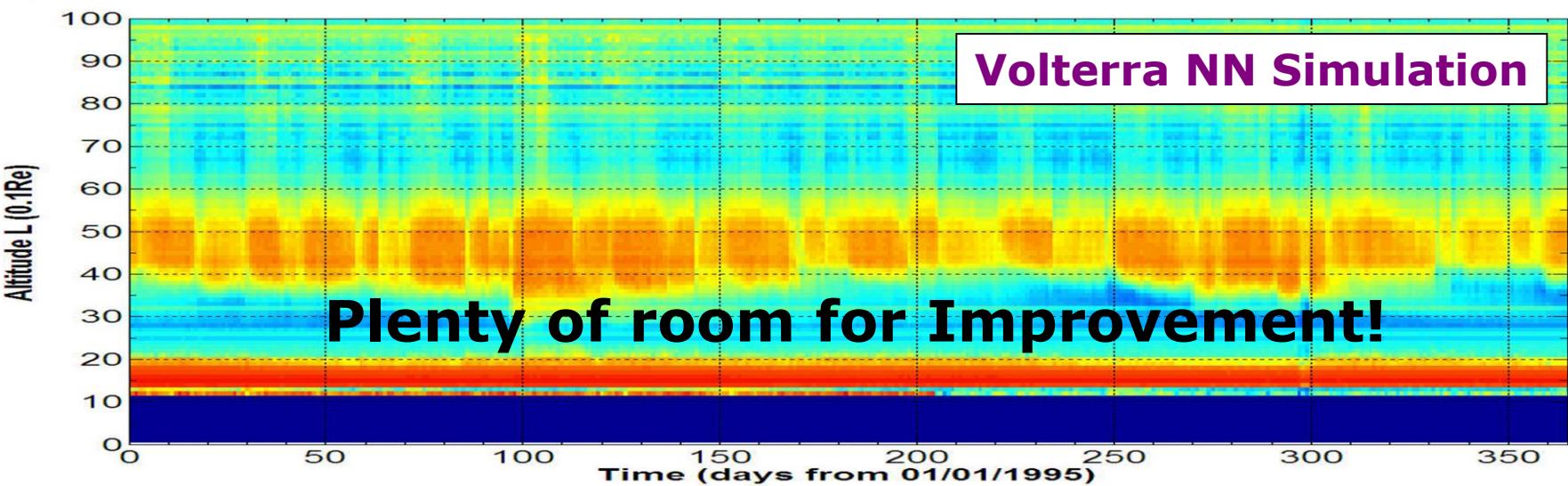


# NAR(23)-Model (1995)

a)

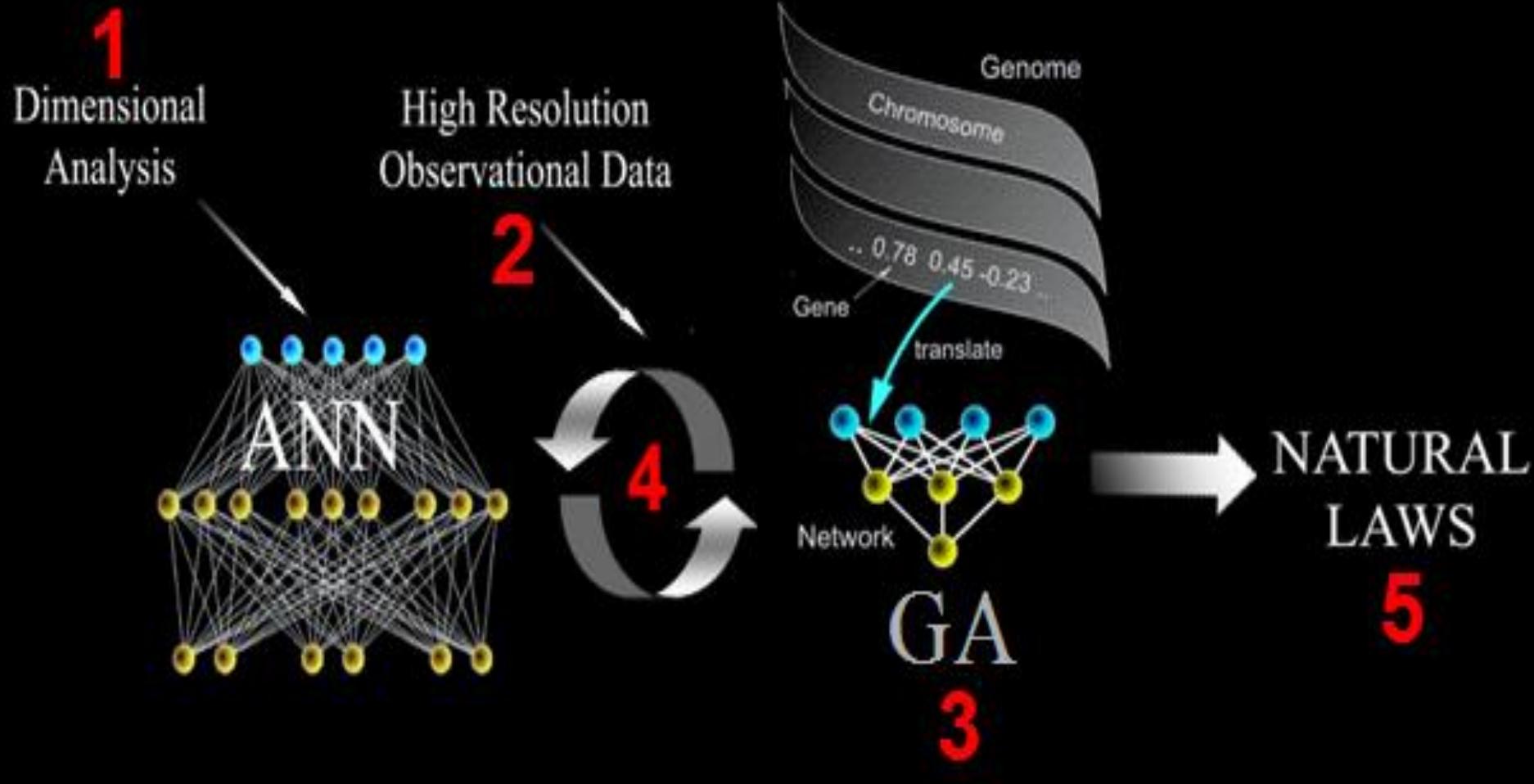


b)



- 1) To improve the simplest NAR model**
- 2) To include SW,IP,MS driver inputs so as to test NARMAX models and to assess the role of the drivers**
- 3) To compare with state-space models**
- 4) To implement real-time updating of the data and re-training of best network models**
- 5) To work towards an online “monitor”**

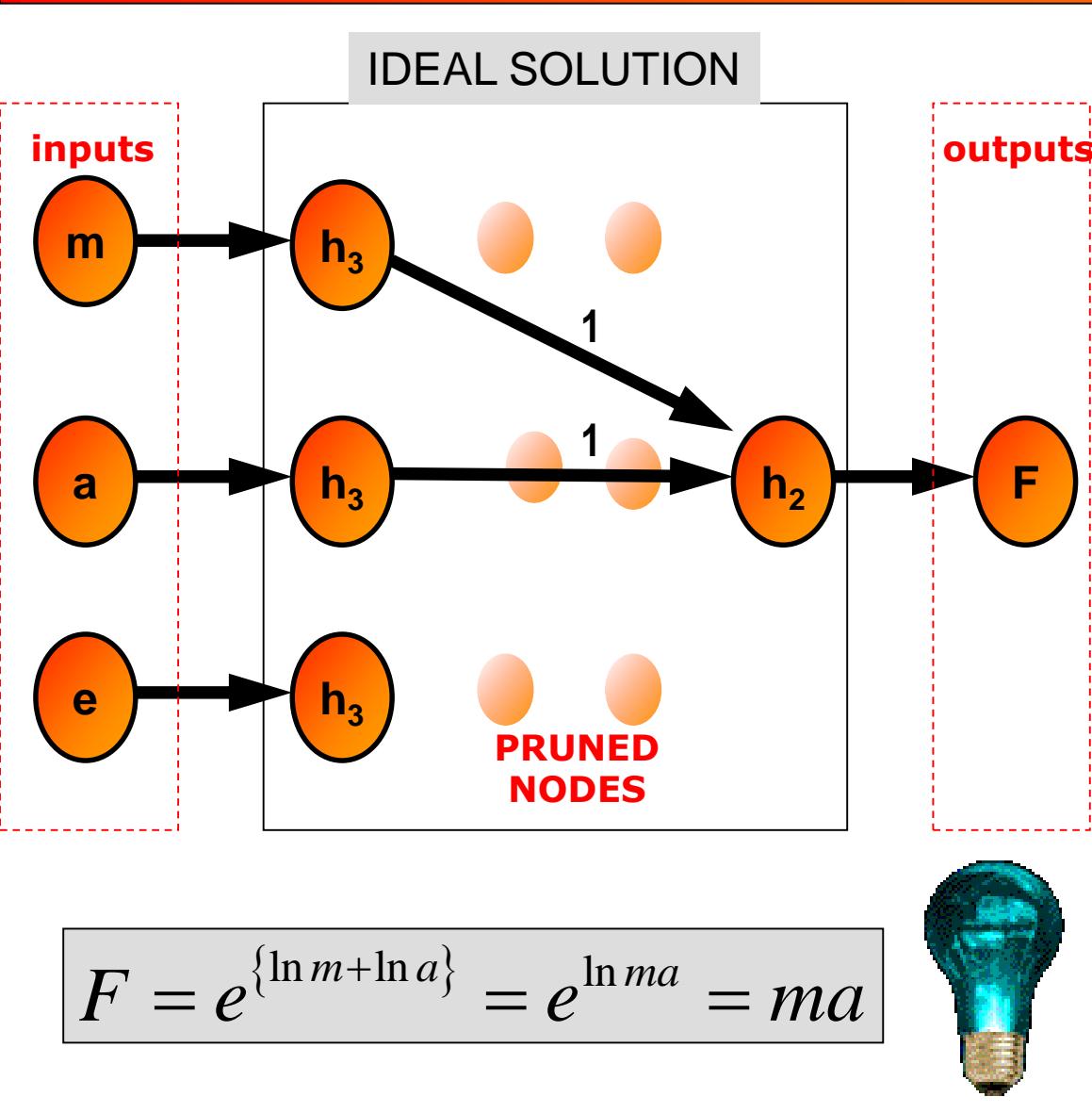
# Strategy 4: Solving the permutation problem with evolutionary networks



Successfully applied to HII galaxies abundance law deduction

[Taylor M and Diaz AI (2007) PASP 374:104-110]

# How it Works: e.g. Newton's Law



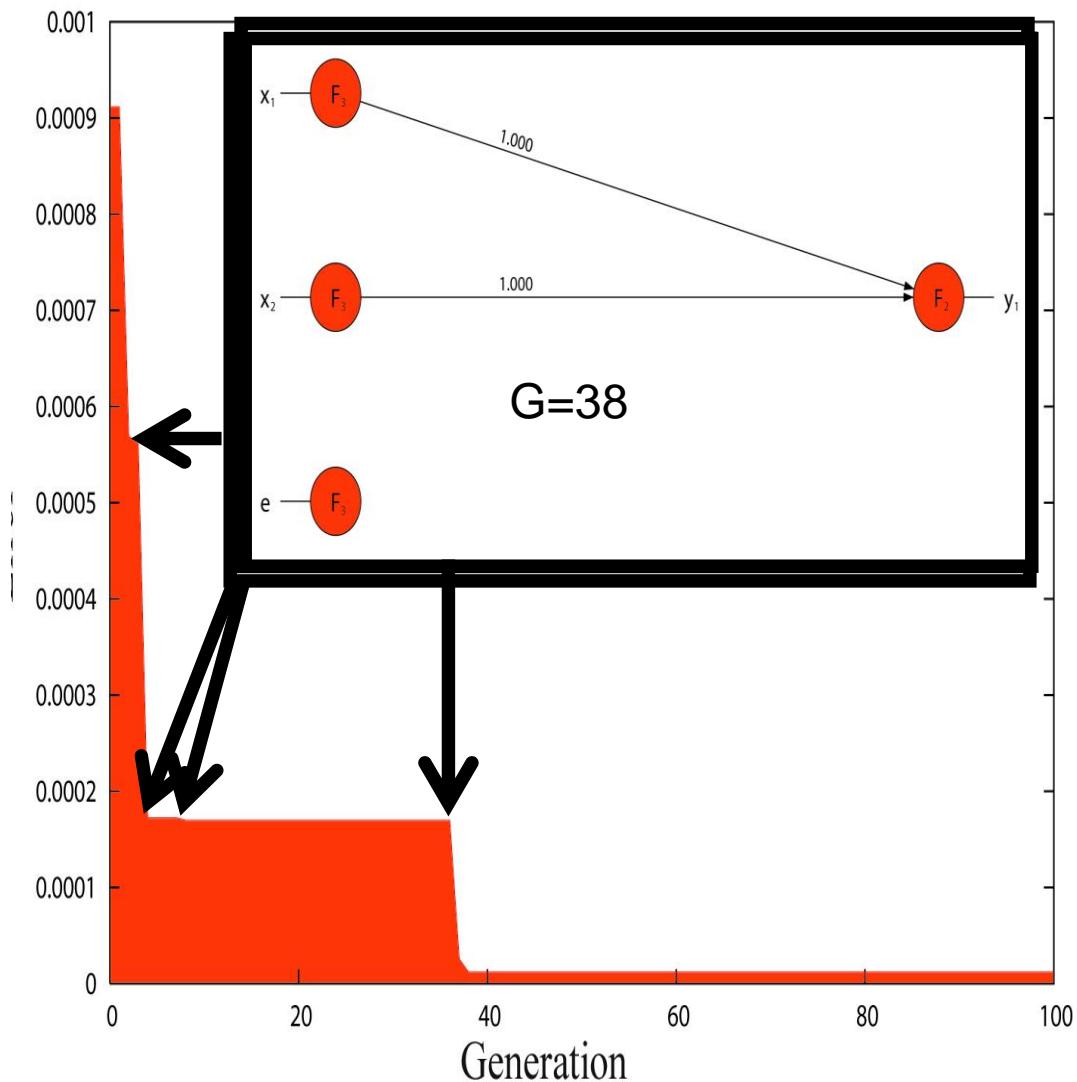
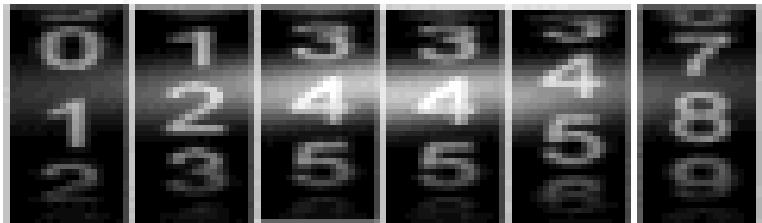
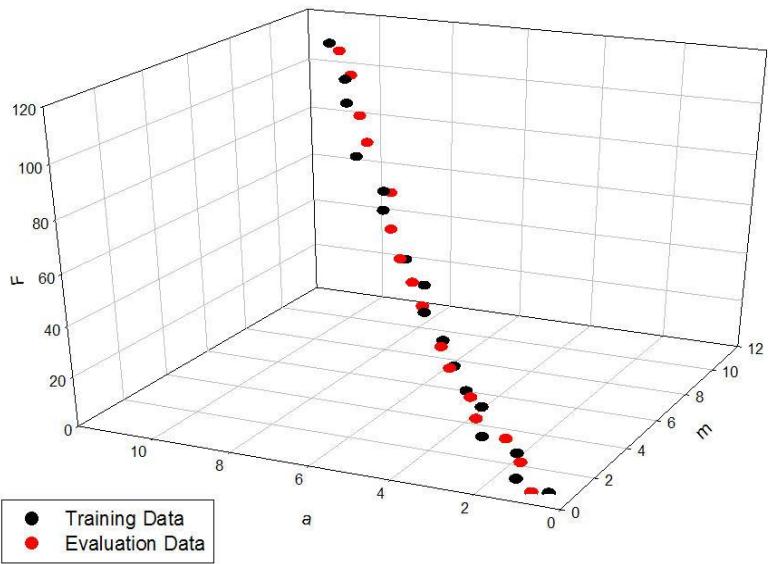
## Neural Node Functions

|             |   |                              |                    |
|-------------|---|------------------------------|--------------------|
| $h(s)_1$    | = | $s$                          | Identities         |
| $h(s)_2$    | = | $e^s$                        | Power Laws         |
| $h(s)_3$    | = | $\ln  s $                    |                    |
| $h(s)_4$    | = | $1/(1 + e^{-s})$             | Sigmoids           |
| $h(s)_5$    | = | $(1 + e^{-s})/2(1 + e^{-s})$ |                    |
| $h(s)_6$    | = | $\tanh s$                    |                    |
| $h(s)_7$    | = | $\sin s$                     | Harmonic Functions |
| $h(s)_8$    | = | $\cos s$                     |                    |
| $h(s)_9$    | = | $e^{-s^2}$                   | Gaussian           |
| $h(s)_{10}$ | = | 1                            | Constants          |

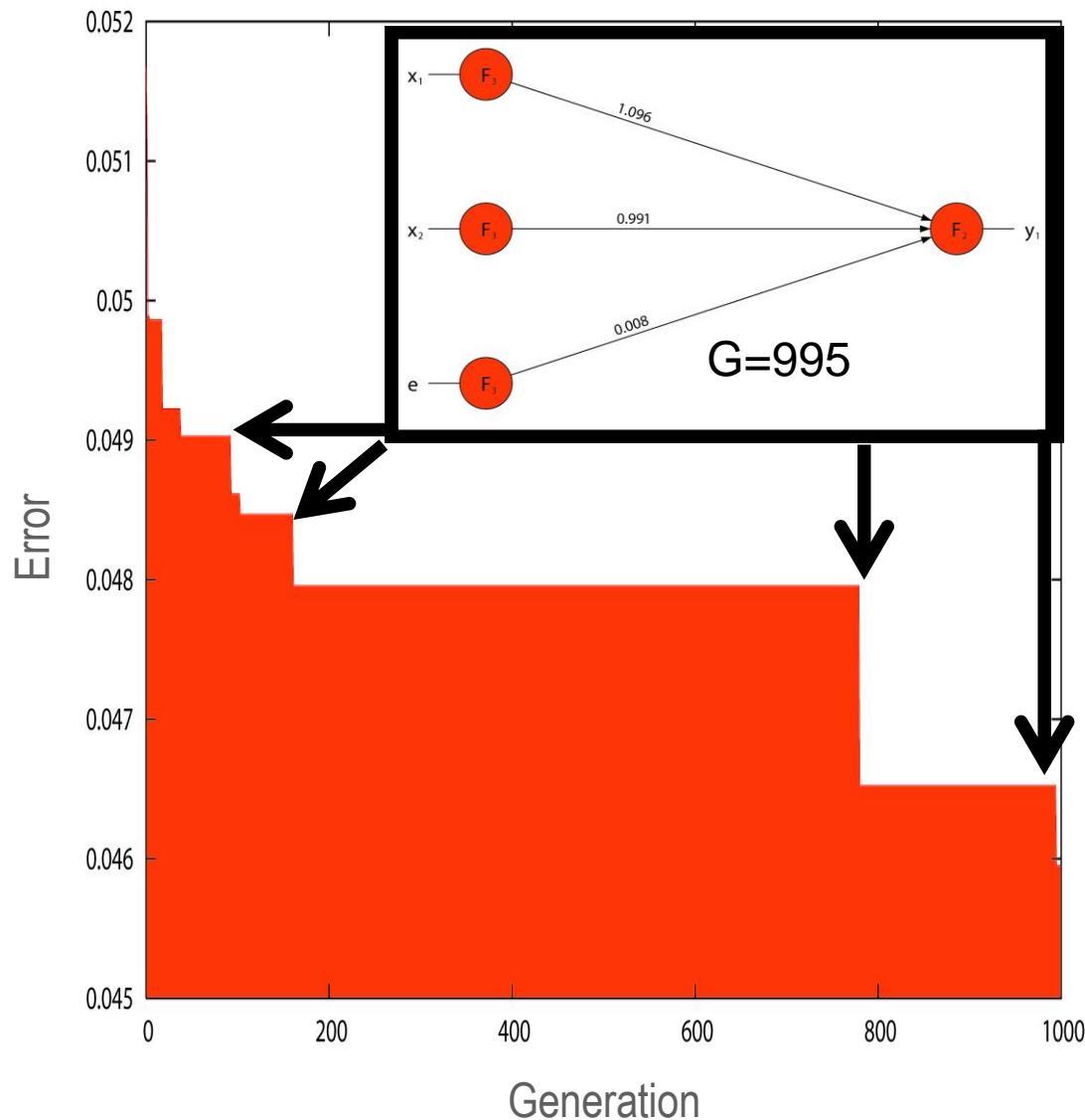
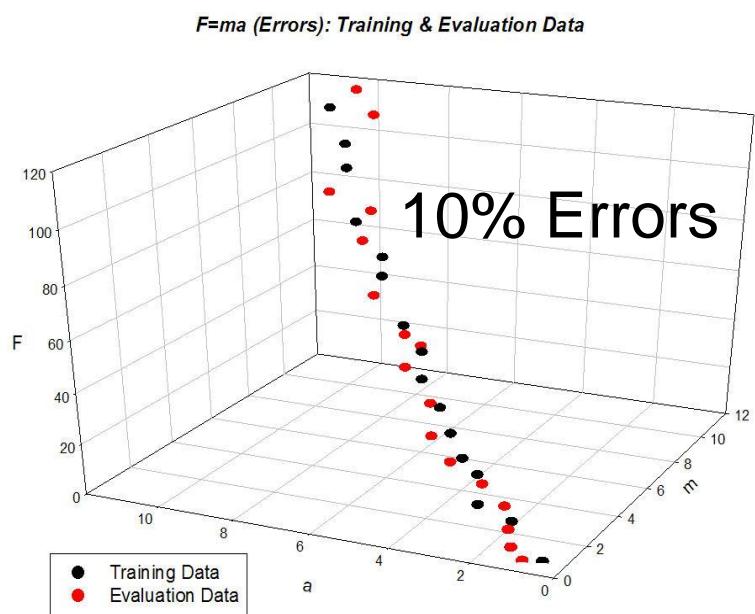


# Newton's Law: $F=ma$ (EXACT)

$F=ma$  (EXACT): Training & Evaluation Data



# Newton's Law: $F=ma$ (10% Errors)



- 1) To use the results of dimensional analysis in Strategy 2 to evolve dimensionless nets**
- 2) To get mainframe CPU time to do this massive run on the 10 year, 11 parameter dataset (or a sizeable chunk of it)**
- 3) To use the results to deduce the functional relation “f” between  $J_e$  and the driver inputs that include time-dependence**

# "The Armageddon Storm" (2012)?

Solar Flare Threatens The Earth With a Storm  
New York Times (1857-Current file) Jul 16, 2000; ProQuest  
pg. 21

## *Solar Flare Threatens The Earth With a Storm*

WASHINGTON, July 15 (AP) — A huge solar eruption has taken place as predicted, the National Oceanic and Atmospheric Administration said, bringing a possibility of disrupted radio transmissions and bright northern lights.

The solar flare, which took place about 6:24 a.m., ejected billions of tons of plasma and charged particles into space, some of it heading toward Earth at three million miles per hour. The mass ejection then headed for the Earth's magnetic field.

The sun is now in the most intense phase of its 11-year sunspot cycle.

S  
C  
E  
N  
E



How well can machines be trained to predict space weather storms?

Michael Taylor  
michael@space.noa.gr

**“Living with a star” in fear?**

(with good net forecasts, hopefully not!)



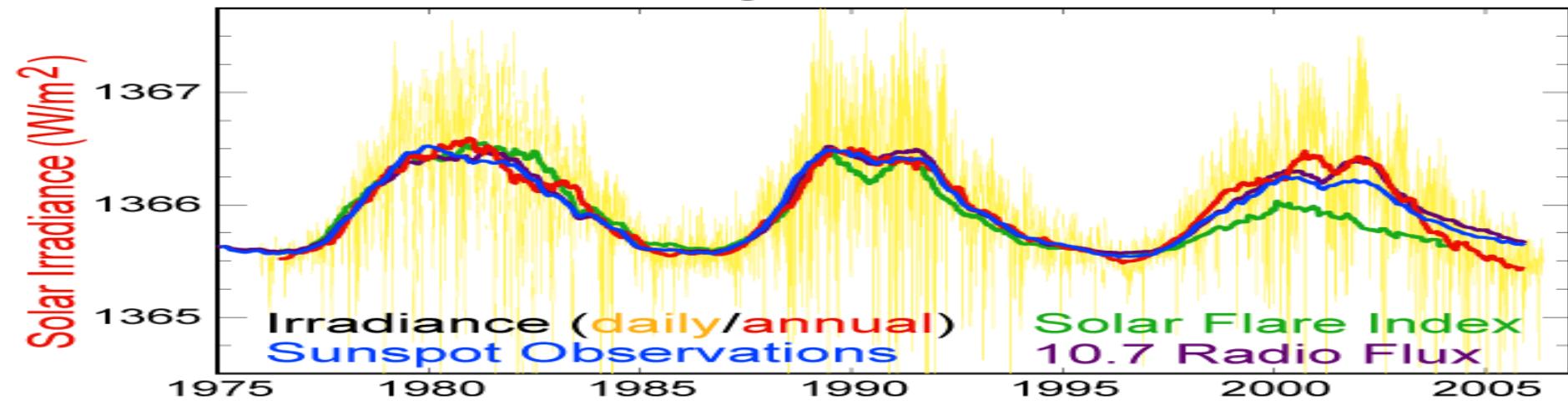
*Ευχαριστώ ολους στο ISARS για την φιλοχενία σας*

**Many thanks**

# Solar Cycle and Sunspot Predictions

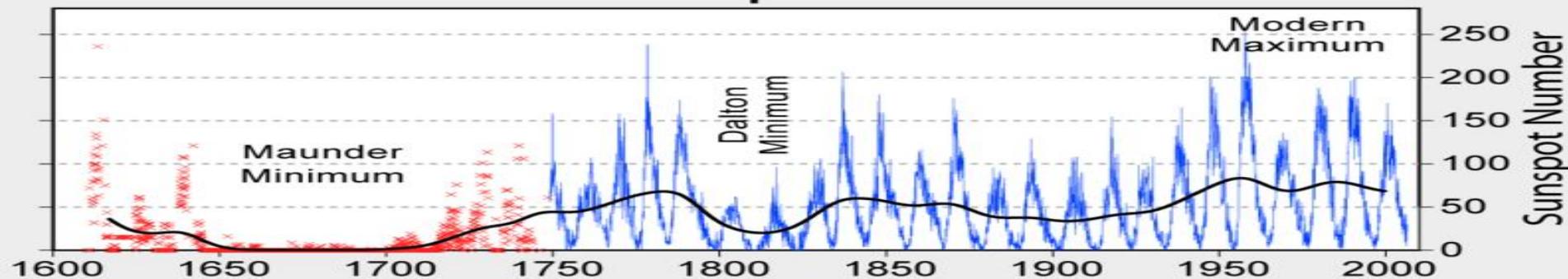


## Solar Cycle Variations



[Plotted from data compiled by: Willson RC and Mordvinov AV (2003). GRL 30(5):1199]

## 400 Years of Sunspot Observations



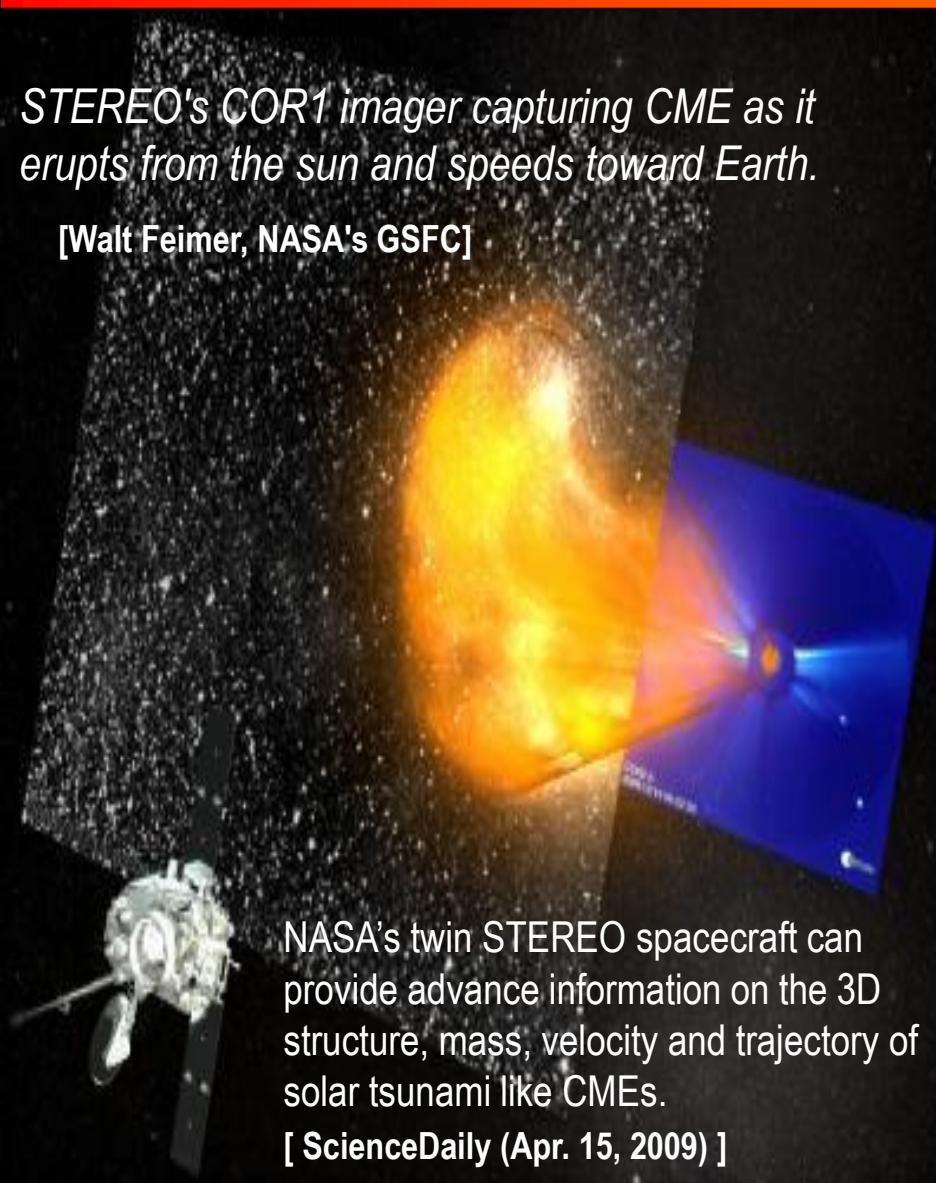
[Plotted from data compiled by: Hoyt DV and Schatten KH (1998a,b) Solar Physics 179:189-219 and 181:491-512]

# Coronal Mass Ejections (CMEs)

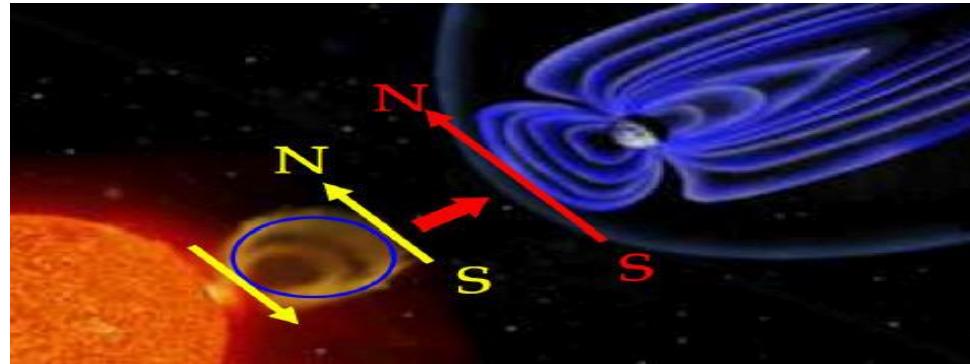


*STEREO's COR1 imager capturing CME as it erupts from the sun and speeds toward Earth.*

[Walt Feimer, NASA's GSFC]

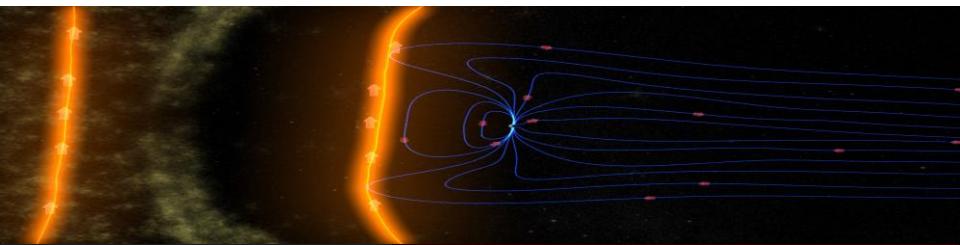


The sun changes polarity in each cycle. In the maximum of Cycle 25 in 2012 it will eject mostly N-S CMEs.



THEMIS observed that a 10 times larger than expected breach in Earth's magnetic shield occurred due to the alignment of the solar plasma and Earth magnetic fields on the sunward side.

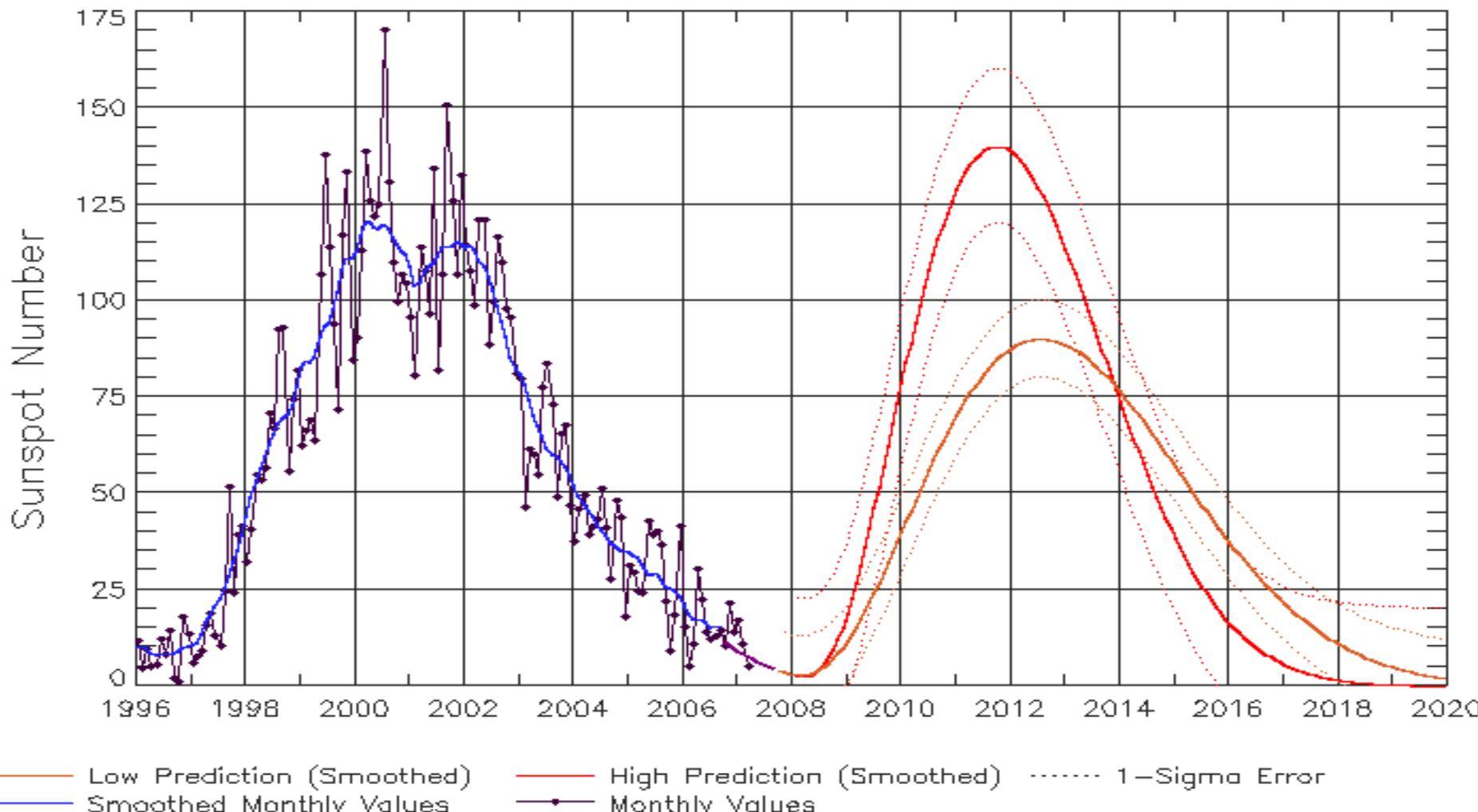
Magnetic storms during the solar cycle maximum could now be stronger than the storms of the previous cycle.



# SSN Prediction Cycle 25



Solar Cycle 24 Sunspot Number Prediction  
Data Through 31 Mar 07



Low Prediction (Smoothed)  
Smoothed Monthly Values

High Prediction (Smoothed)  
Monthly Values

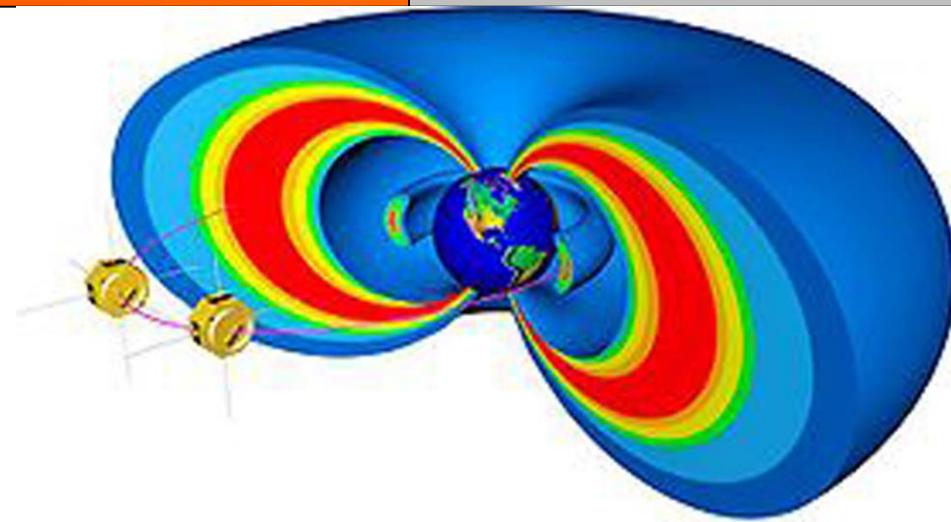
Updated 2007 Apr 20

NOAA/SEC Boulder, CO USA

How well can machines be trained to predict space weather storms?

Michael Taylor  
michael@space.noaa.gr

# NASA's RBSPs



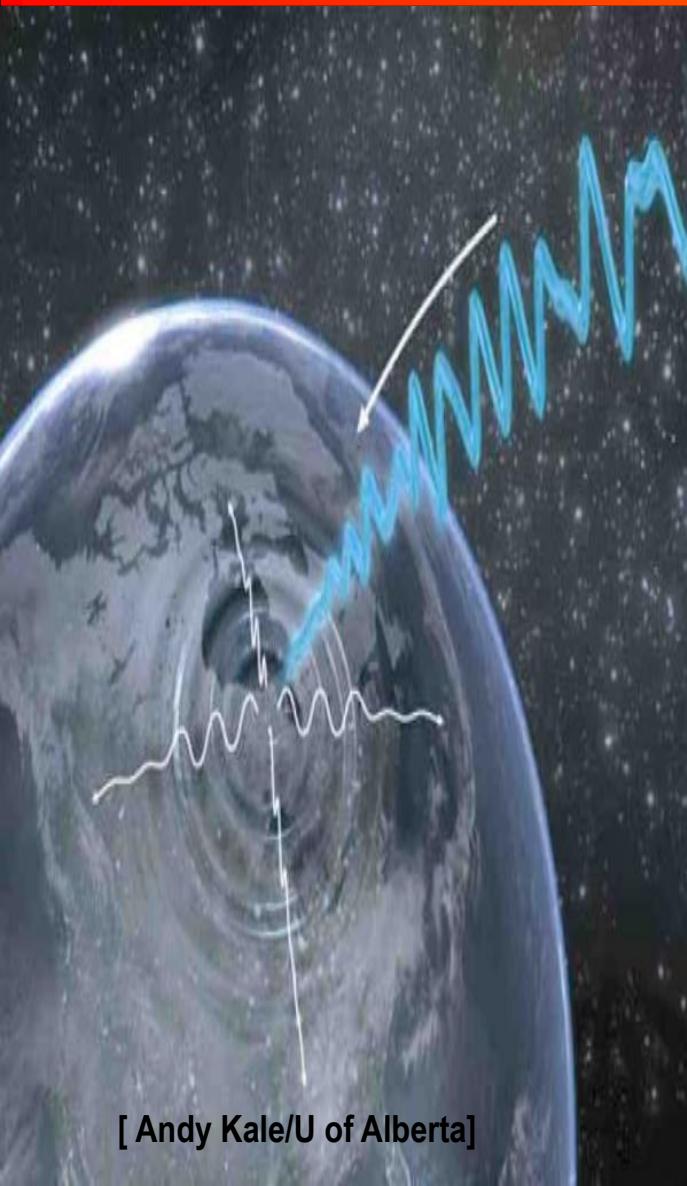
The Radiation Belt Storm Probes (RBSP) is a NASA mission under the GSFC's "Living With a Star" program.

The mission of RBSP is to gain scientific understanding (to the point of predictability) of how populations of relativistic electrons and ions in space form or change in response to changes in solar activity and the solar wind.

The RBSP mission is currently scheduled for 2011 and will last 2 years.

On 16 March 2009 NASA awarded the United Launch Alliance (ULA) a contract to launch RBSP using an Atlas V 401 rocket

# Geomagnetic Sub-Storm Epicentres



[ Andy Kale/U of Alberta]

Magnetic blast waves can be used to pinpoint and predict the location where space storms dissipate their massive amounts of energy. These storms can dump the equivalent of 50 gigawatts of power, or the output of 10 of the world's largest power stations, into Earth's atmosphere.

University of Alberta research team uses ground-based observatories spread across northern Canada and the five satellites of the THEMIS mission to detect magnetic disturbances as storms crash into the atmosphere using "space seismology," they look for the eye of the storm hundreds of thousands of miles above Earth. "Undulating auroral features and ripples in Earth's magnetic field began at the same time and propagated away from Sanikulaq, Nunavut, Canada at speeds on the order of 60,000 miles per hour, much like the blast wave from a gigantic explosion," said David Sibeck.

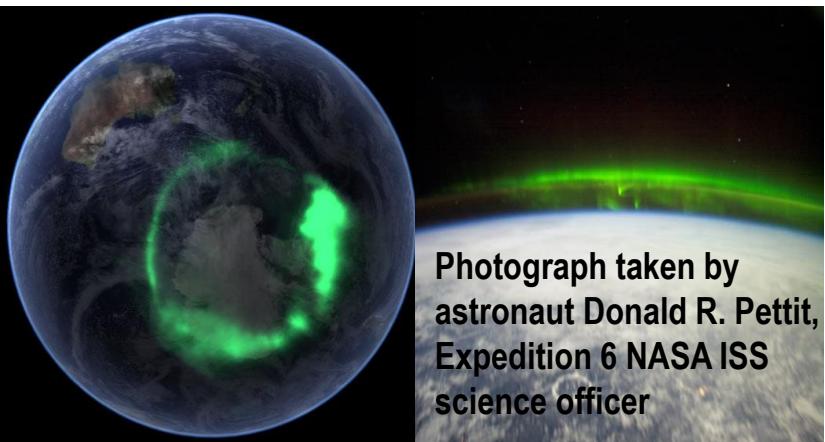
Guided by Earth's magnetic field, the magnetic tremors rocket through space toward Earth. These geomagnetic substORMS trigger magnetic sensors on the ground as they impact the atmosphere. The effects of these storms, and the most spectacular displays of the Northern Lights, follow a few minutes later.

[ ScienceDaily (May 29, 2009) ]

# Aurorae & THEMIS

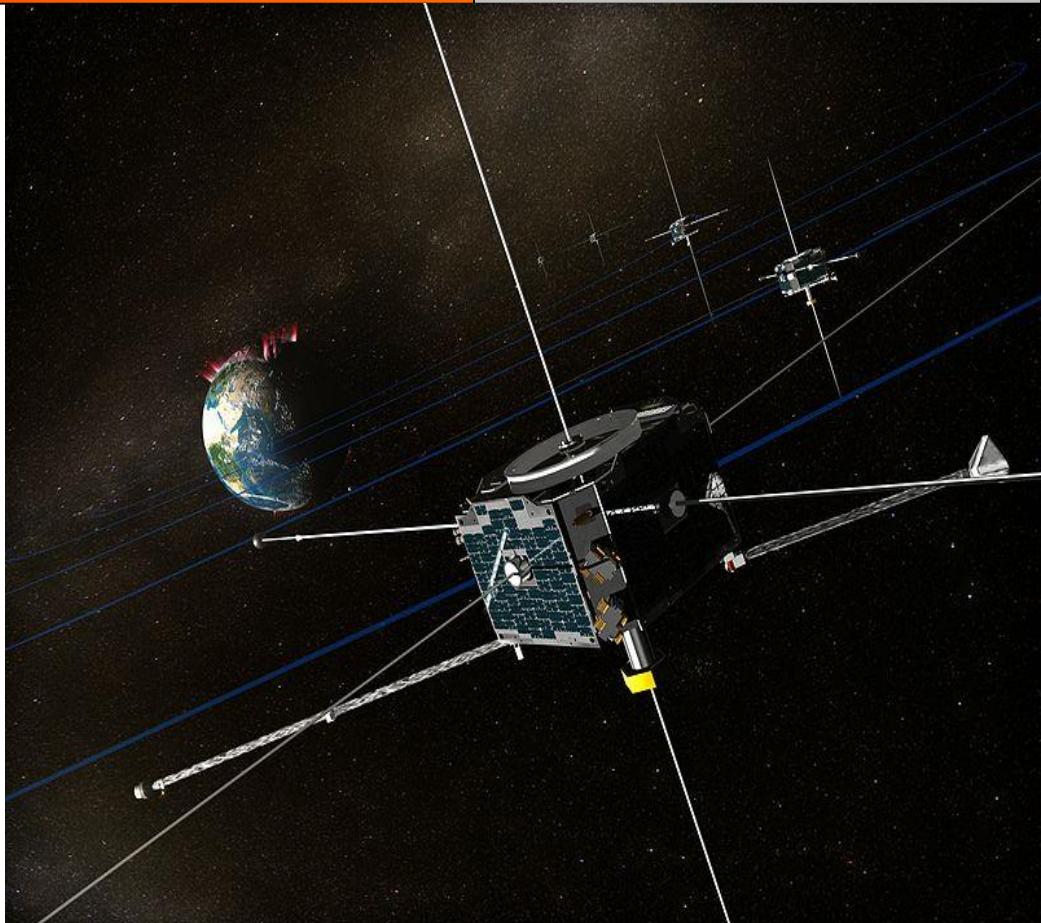


**Aurora Borealis** captured by NASA's FAST satellite  
(September 11, 2005)



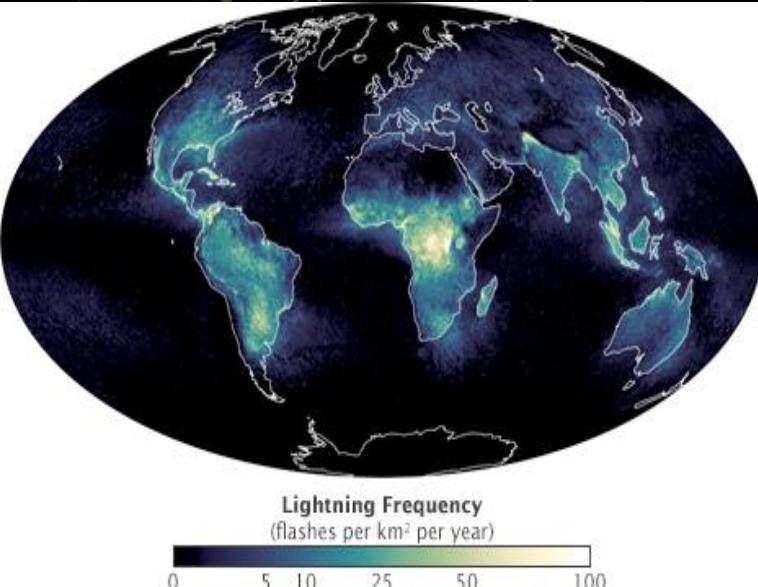
Photograph taken by  
astronaut Donald R. Pettit,  
Expedition 6 NASA ISS  
science officer

**Aurora Australis** captured by NASA's IMAGE satellite  
(September 11, 2005) and overlaid onto NASA's satellite-based Blue Marble image.



The THEMIS mission uses a constellation of 5 NASA satellites to study energy releases from Earth's magnetosphere known as substorms

# The Safe-Zone Lightning Connection



The flash we see from lightning is just part of the total radiation it produces. Lightning also generates radio waves. These radio waves are bent by electrically charged gas trapped in the Earth's magnetic field. That causes the waves to flow out into space along the Earth's magnetic field lines.

According to the lightning theory, radio waves clear the safe zone by interacting with the radiation belt particles, removing a little of their energy and changing their direction. This lowers the mirror point, the place above the polar regions where the particles bounce. Eventually, the mirror point becomes so low; it is in the Earth's atmosphere. When this happens, the radiation belt particles can no longer bounce back into space, because they collide with atmospheric particles and dissipate their energy.

Radio wave data from the Radio Plasma Imager on the Imager for Magnetopause to Aurora Global Exploration (IMAGE) spacecraft, combined with archival data from the Dynamics Explorer spacecraft, show that radio wave activity in the safe zone closely follows terrestrial lightning patterns observed by Micro Lab 1. When magnetic storms, caused by violent solar activity, inject a new supply of high-speed particles into the safe zone, lightning clears them away in a few days.

[ ScienceDaily (Mar. 10, 2005) ]