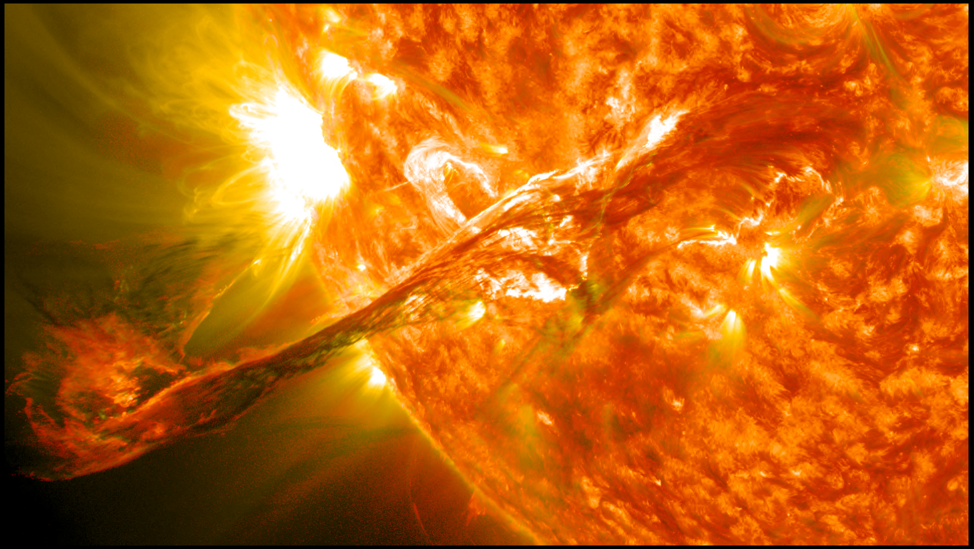
Machine Learning for Space Weather

Enrico Camporeale

Multiscale Dynamics



www.mlspaceweather.org

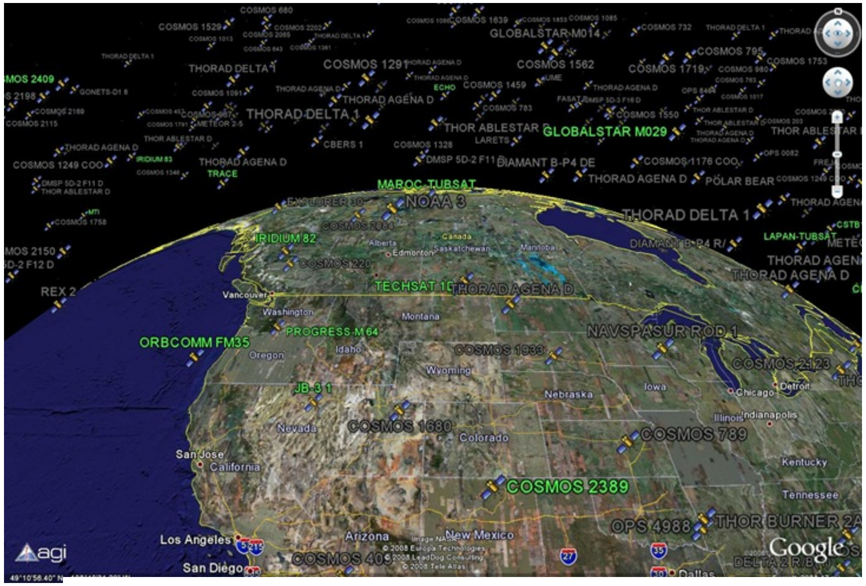
CWI
Centrum Wiskunde & Informatica

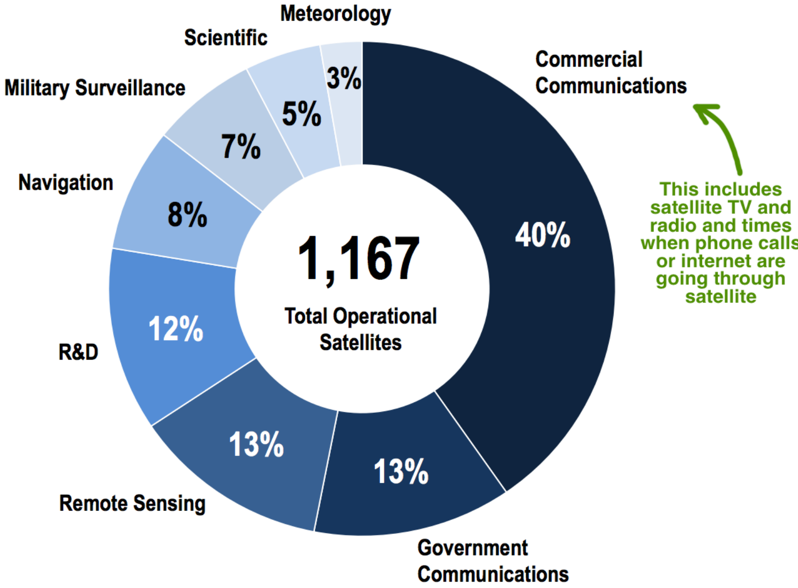
CWI Scientific Meeting
14th October 2016

Imagine a day...

...when satellites

stop working

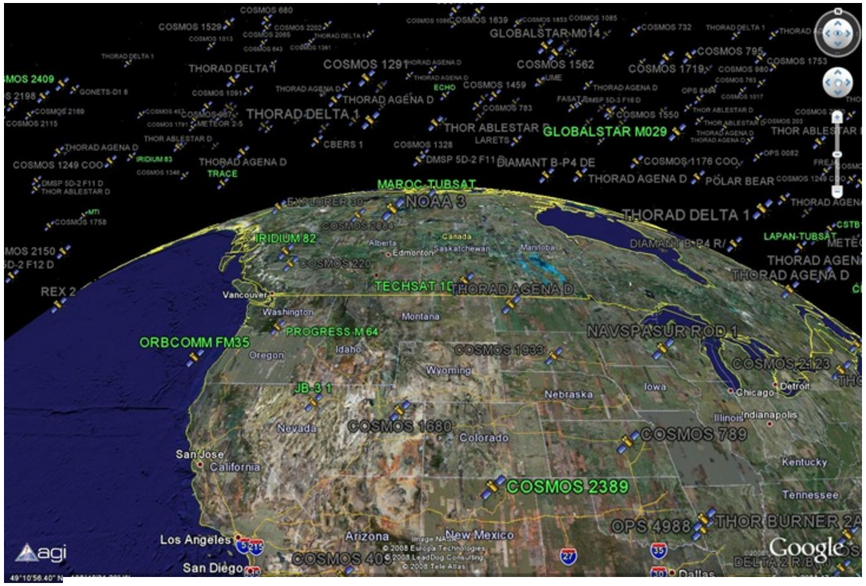


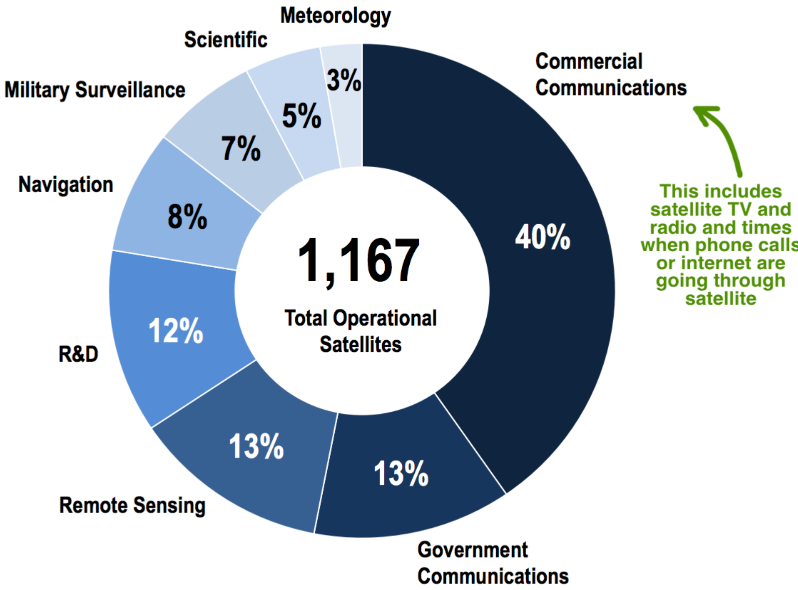


Imagine a day...

...when satellites stop working

- No telecommunications



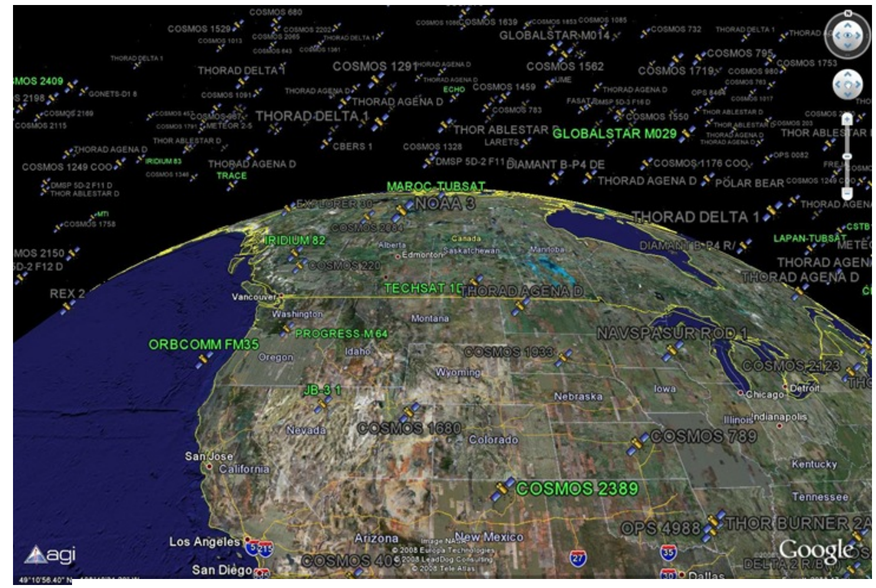


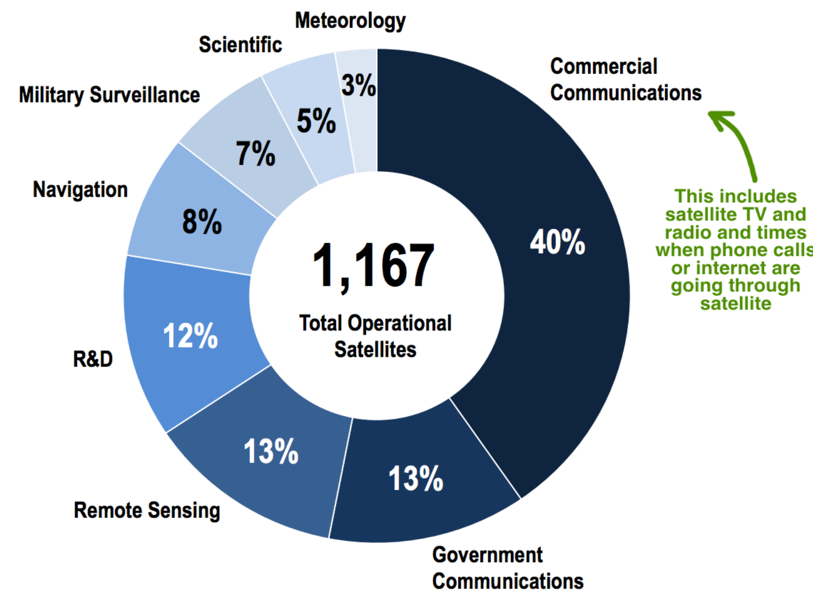
Imagine a day...

...when satellites

stop working

- No telecommunications
- Terrestrial and undersea lines overloaded



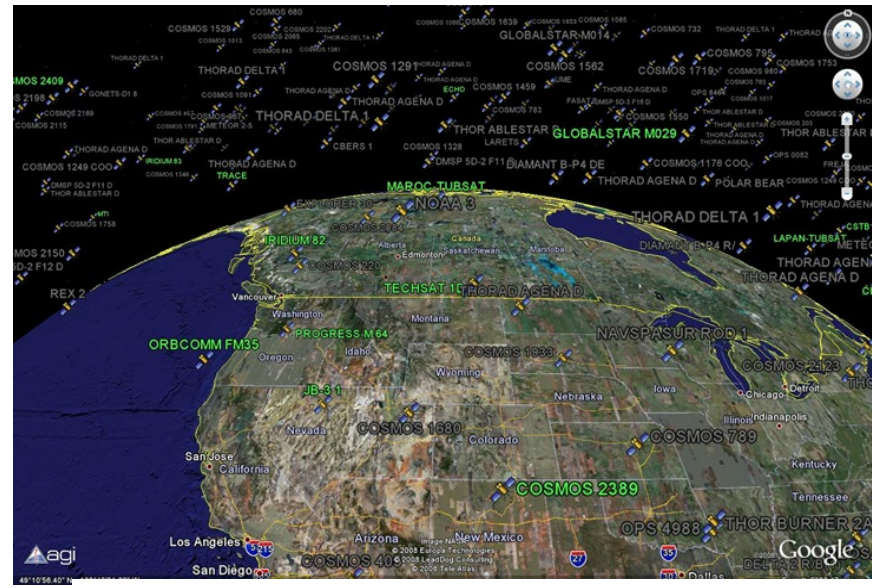


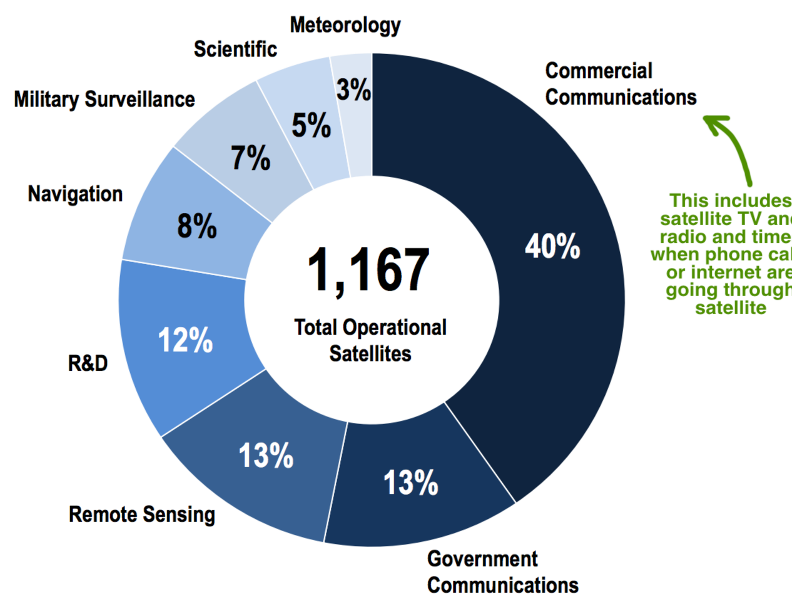
Imagine a day...

...when satellites

stop working

- No telecommunications
- Terrestrial and undersea lines overloaded
- No military surveillance
- No weather forecast





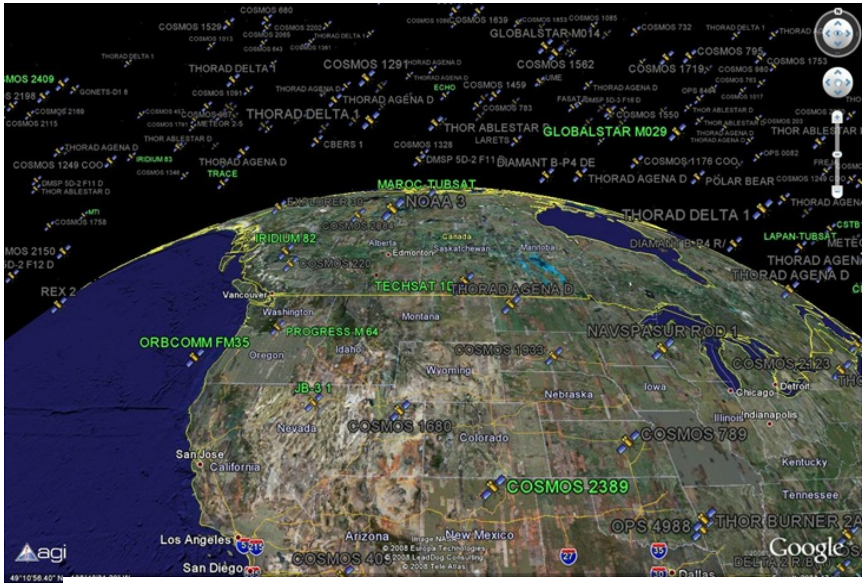
Imagine a day...

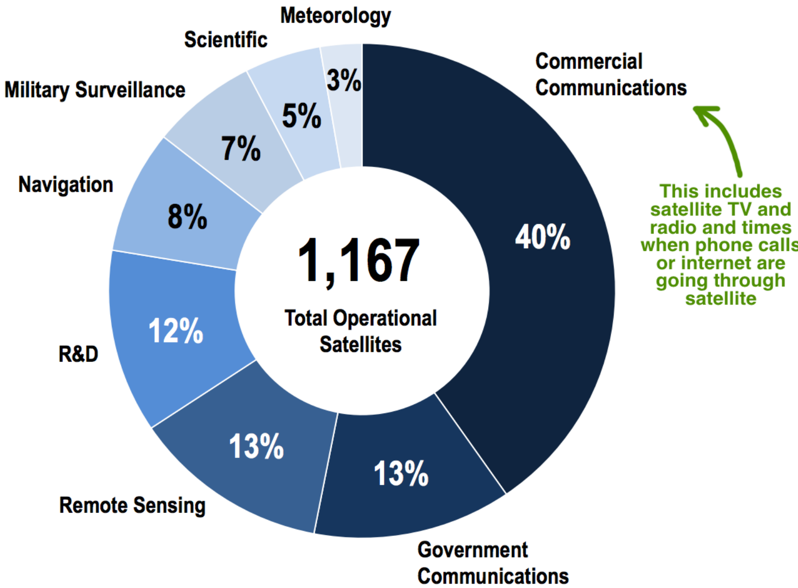
...when satellites

stop working

- No telecommunications
- Terrestrial and undersea lines overloaded
- No military surveillance
- No weather forecast

- **No GPS**





Imagine a day...

No GPS

- Drilling, agriculture
- Aviation
- Railways

Imagine a day...

No GPS

GPS provides precise time stamps to:

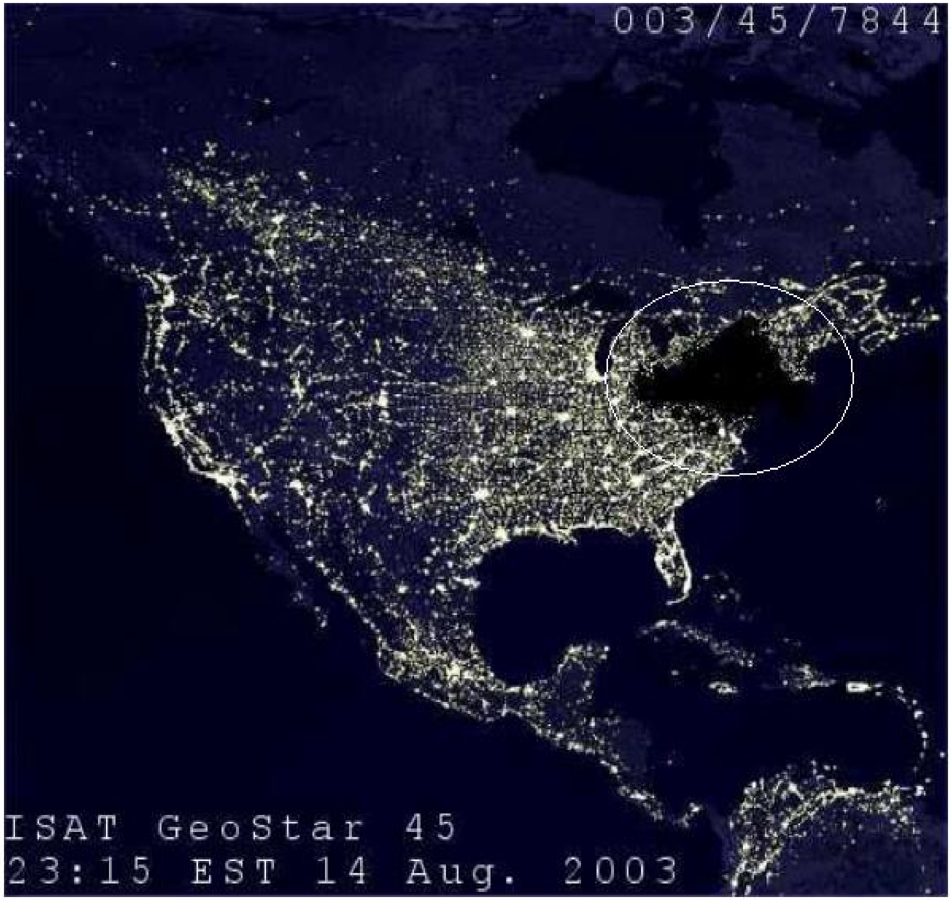
- Communications systems
- Electrical power grids
- Stock market

...and this is only the start

Imagine if...

- Several high-voltage transformers would be damaged, causing **widespread blackouts**, affecting:

- Transportation
- Water supply
- Communication
- Industry



...and finally...

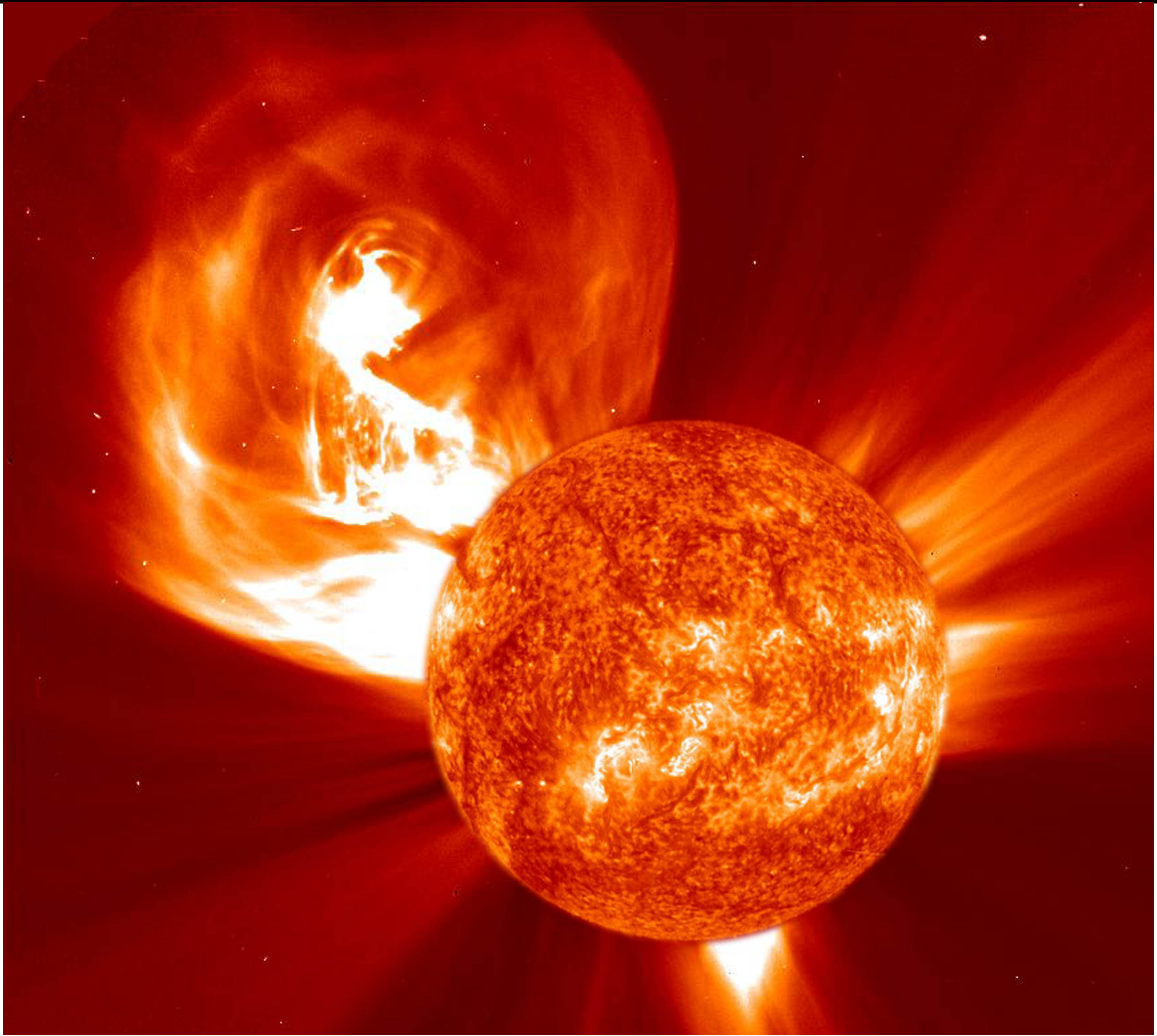
Because military satellites are blinded the USA lost their ability of detecting long-range ballistic missiles

...and finally...

- Because military satellites are blinded the USA lost their ability of detecting long-range ballistic missiles
- President Trump decides to launch a pre-emptive nuclear strike on North Korea

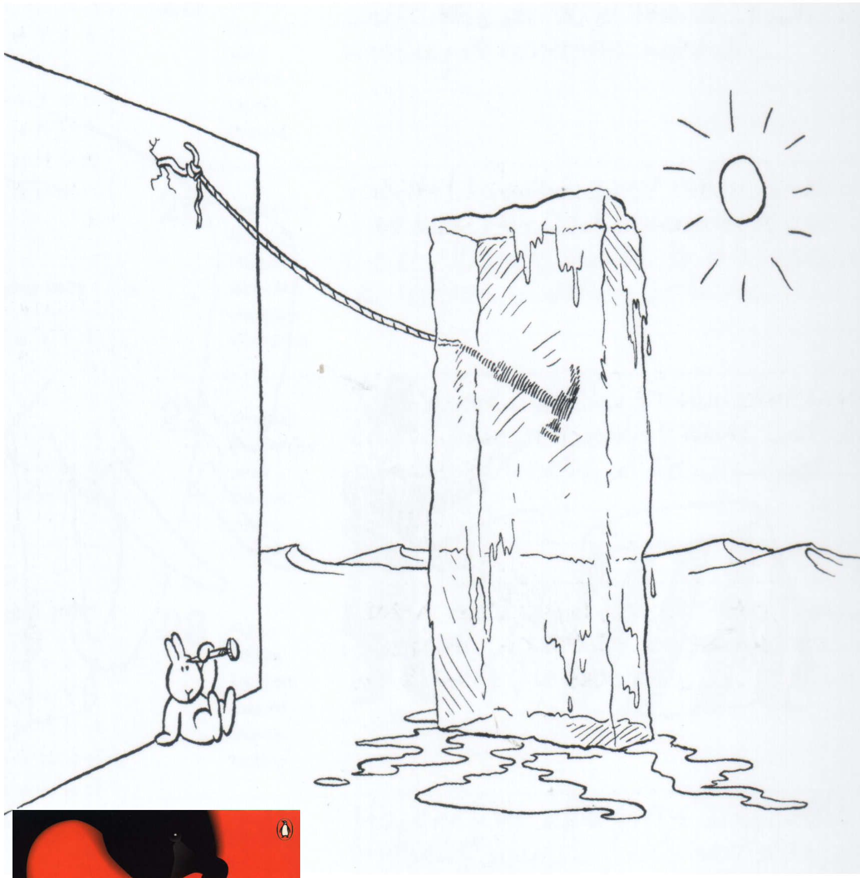
What can possibly cause such a disaster?

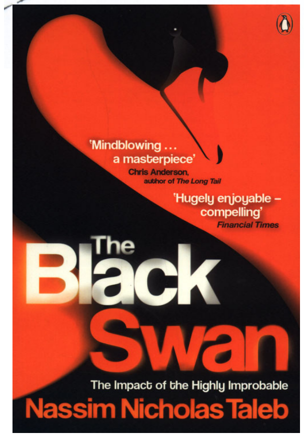
What can possibly cause such a disaster?



A disaster waiting to happen

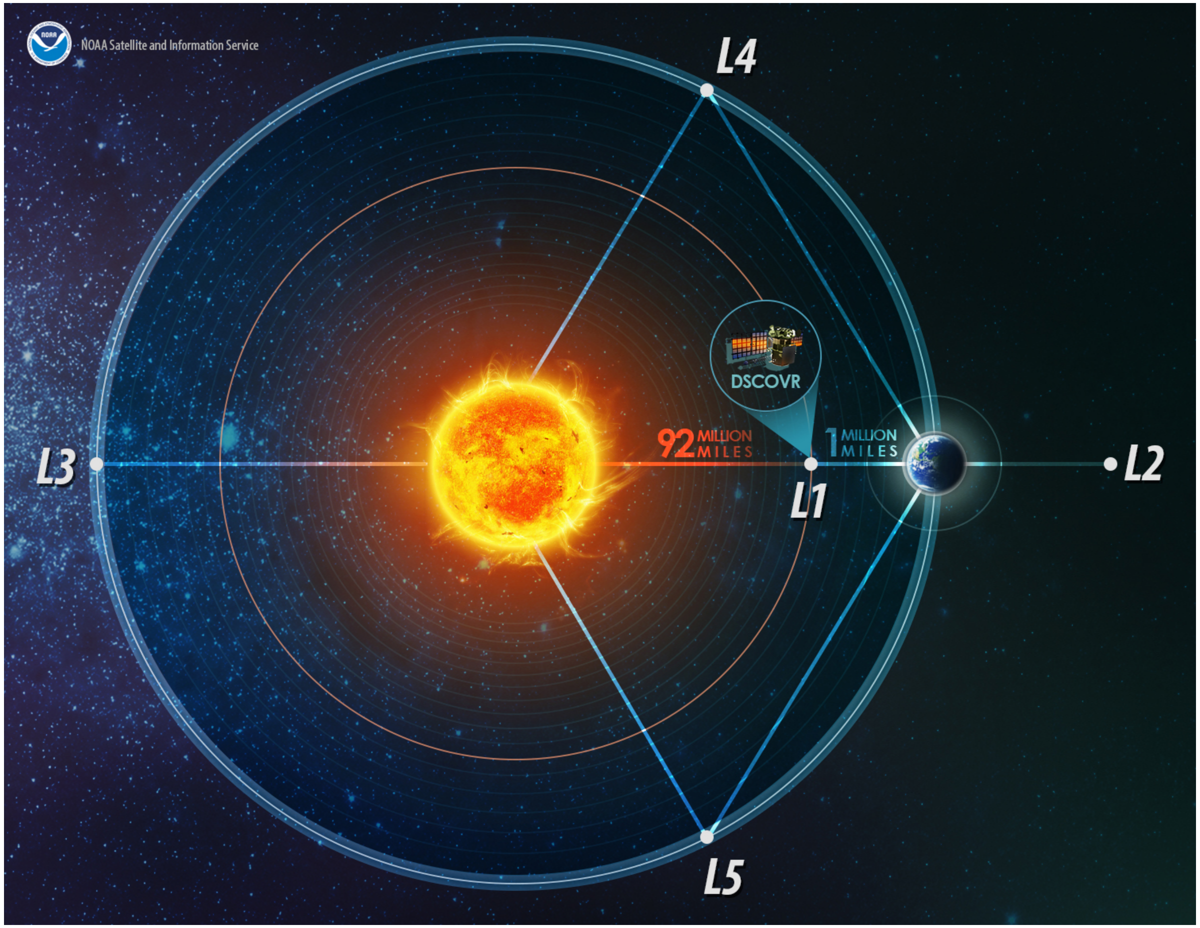
- The adverse impact of space weather is estimated to cost **$200–$400 million** per year;
- losses to satellite companies range from thousands of dollars for temporary data outages up to **$200 million** to replace a satellite;
- economists also estimate that timely warnings of geomagnetic storms to the electric power industry would save approximately **$150 million** per year;
- a 1% gain in continuity and availability of GPS would be worth **$180 million** per year.
- **a “big one” would cause $2.6 trillion damage**





From
The book of bunny suicides

Space Weather predictions



The 1st Lagrangian point is our privileged Sun observatory from which we gather:

– remote images

– in situ data

L1 data are then complemented with ground-based and low altitude satellites data

Our project (CWI / INRIA)

Machine Learning for Space Weather

Home Team Projects Publications Resources Jobs Contact Us

Coupling physics-based simulations with Artificial Intelligence

Goal

In this project we aim at enhancing the current state-of-the-art simulations for Space Weather, by using prior knowledge gathered from historical satellite data. Several Machine Learning techniques will be used for data-mining, classification, and regression. The long-term objective of the project is the creation of a portfolio of data-enhanced reduced models, along with automated rules for model selection. Depending on the real-time conditions observed by satellites, the resulting 'grey-box' model should choose the relative importance between physical and empirical estimations.

CWI/INRIA consortium

This project has started as funded by a CWI/INRIA collaboration. However, several other parties have joined in this activity, either as external collaborators, or more actively involved. Visit the team page for a list of all the people involved.

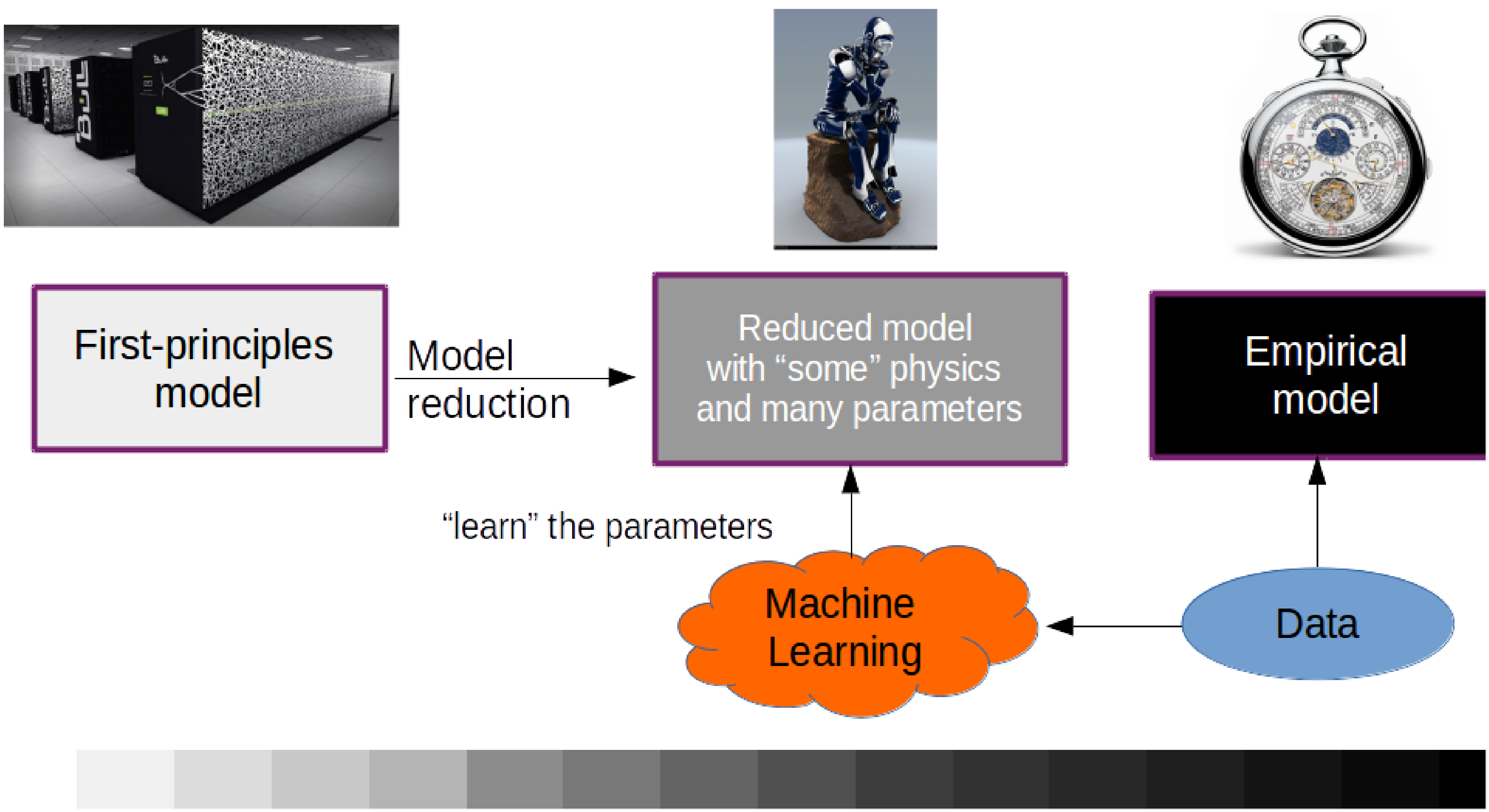
CWI Centrum Wiskunde & Informatica

Inria INVENTEURS DU MONDE NUMÉRIQUE

CWI is the Dutch National Center for Mathematics and Computer Science. INRIA is the French Institute for Research in Computer Science and Automation.

www.mlspaceweather.org

The gray-box paradigm



A quick overview

- Information theory
  - Wing, S., Johnson, J. R., Camporeale, E., & Reeves, G. D. (2016). Information theoretical approach to discovering solar wind drivers of the outer radiation belt. *J. Geophys. Res.*
- Uncertainty Quantification
  - Camporeale, E. , Shprits, Y., Chandorkar, M., Drozdov, A., Wing, S. (2016) On the propagation of uncertainties in radiation belt simulations, *Space Weather*
  - Camporeale, E. , A. Agnihotri, C. Rutjes (2016) Adaptive selection of sampling points for uncertainty quantification, *J. Sci. Comp.*, under review.
- Machine Learning
  - Chandorkar, M., Camporeale, E., Wing, S. (2016) Gaussian Processes Autoregressive Models for Forecasting the Disturbance Storm Time Index, *J. Space Weather Space Clim.*, under review

Information theory

Question: which quantity in the solar wind drives the changes in electron flux in the radiation belt?

Entropy: $H(x) = -\sum_{\aleph_1} p(\hat{x}) \log p(\hat{x}); \quad H(y) = -\sum_{\aleph_2} p(\hat{y}) \log p(\hat{y})$

Joint Entropy: $H(x, y) = -\sum_{\aleph_1 \aleph_2} p(\hat{x}, \hat{y}) \log p(\hat{x}, \hat{y})$

Mutual information (not directional)

$$\mathrm{MI}(x, y) = H(x) + H(y) - H(x, y)$$

Conditional mutual information

$$\mathrm{CMI}(x, y \mid z) = \sum_{\aleph_1 \aleph_2 \aleph_3} p(\hat{x}, \hat{y}, \hat{z}) \log \frac{p(\hat{x}, \hat{y} \mid \hat{z})}{p(\hat{x} \mid \hat{z})\, p(\hat{y} \mid \hat{z})} = H(x, z) + H(y, z) - H(x, y, z) - H(z)$$

Information theory

Conditional mutual information

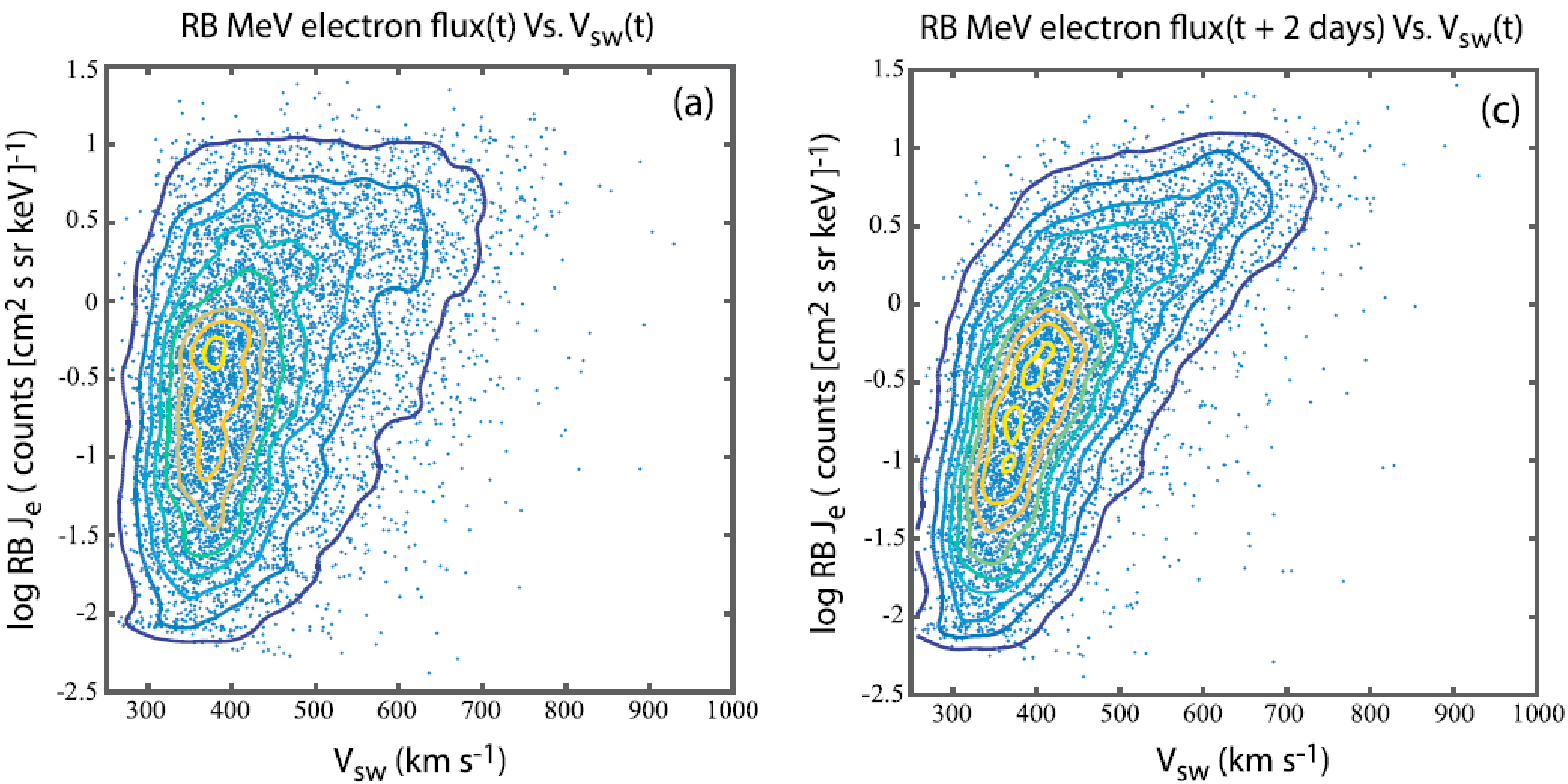
$$\mathrm{CMI}(x,\ y \mid z) = \sum_{\aleph_1 \aleph_2 \aleph_3} p(\hat{x}, \hat{y},\ \hat{z}) \log \frac{p(\hat{x}, \hat{y} \mid \hat{z})}{p(\hat{x} \mid \hat{z})\, p(\hat{y} \mid \hat{z})} = H(x,z) + H(y,z) - H(x,y,z) - H(z)$$

Transfer entropy

$$\mathrm{TE}_{x\to y}(\tau) = \mathrm{CMI}(y(t+\tau), x(t) | yp(t))$$

where $yp(t) = [y(t),\ y(t-\Delta), \ldots, y(t-k\Delta)]$. The transfer entropy can be considered as a conditional mutual information that detects how much average information is contained in an input, x, about the next state of a system, y, that is not contained in the past history, yp, of the system [*Prokopenko et al.*, 2013].

Solar wind velocity vs RB flux



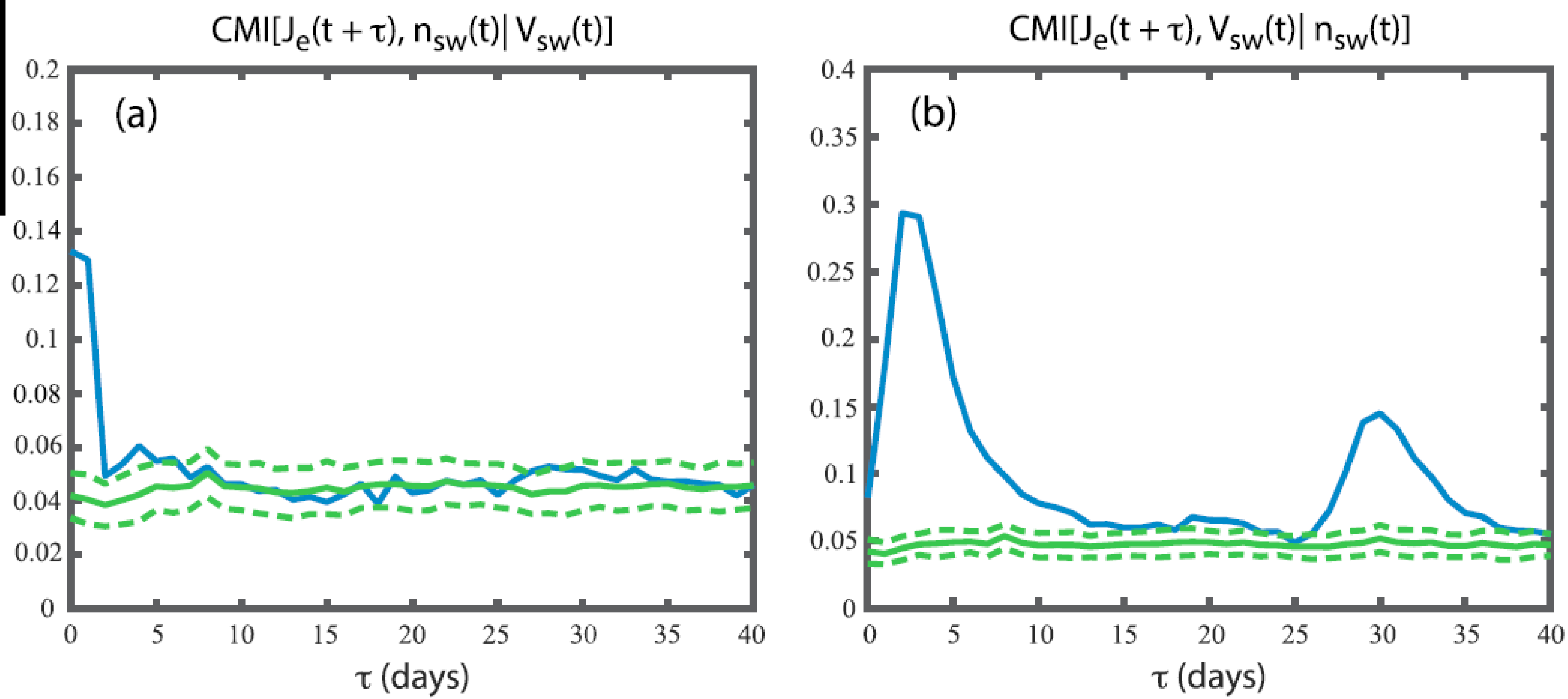
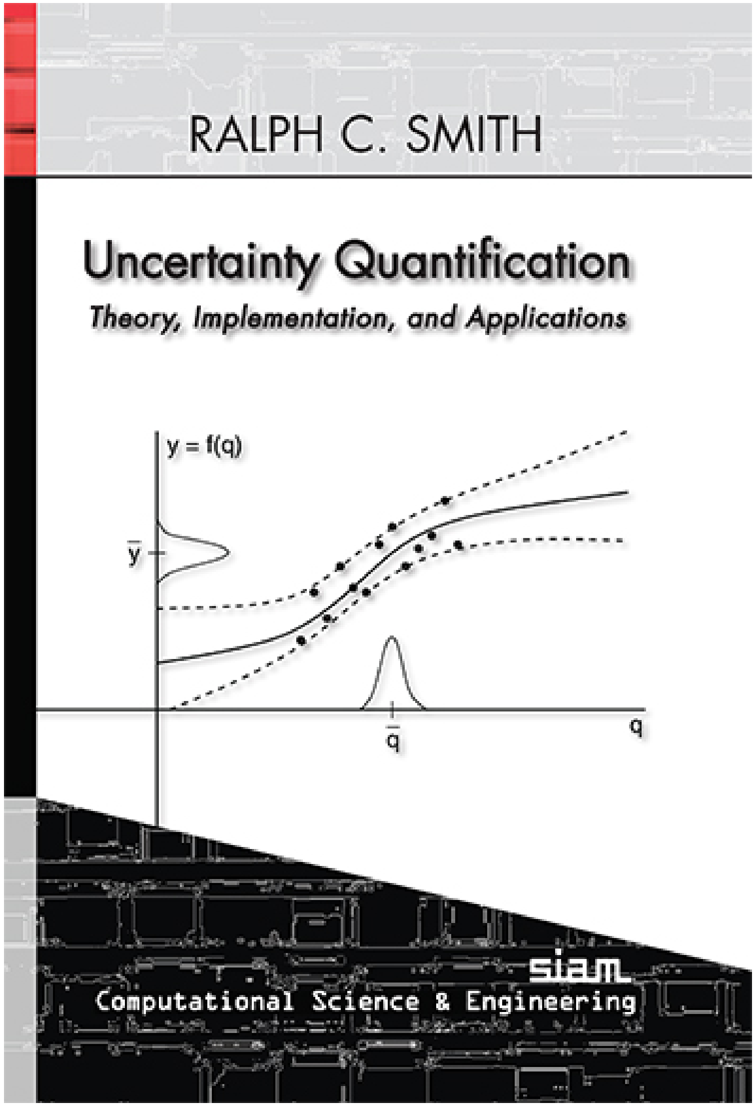


Table 1. Ranking of the Importance of the Solar Wind Parameters Based on Information Transfer to Geosynchronous MeV Electron Flux (J_e) at $\tau_{\max}$, Where $\tau_{\max}$ is the Lag Time When the Information Transfer Peaks[a]

Rank	Solar Wind Parameters	Peak Information Transfer ($it_{\max}$)	Signal-to-Noise Ratio at $\tau_{\max}$	Significance at τ_{max} (σ)	τ_{max} (days)	Prediction Horizon (days)
1	V_{sw}	0.25	6.6	94	2	10[b]
2	IMF $\|\mathbf{B}\|$	0.12	3.9	48	0	2
3	P_{dyn}	0.092	3.4	35	0	2
3	n_{sw}	0.091	3.2	34	0	2
4	σ(IMF B)	0.075	3.9	48	0	2
5	IMF $B_z < 0$	0.064	2.7	26	0	2
6	E_{sw}	0.056	2.9	22	1	5
7	IMF B_y	0.052	2.3	20	0	2
8	IMF $B_z > 0$	0.048	3.1	22	0	2
9	IMF B_x	0.044	2.2	19	0	2

Uncertainty Quantification



Probably, the most used methods are *non-intrusive*, which means that a **black-box computational model** is used repeatedly, for different inputs. The main question is:

What input parameters do you choose? (sampling problem)

- Monte Carlo
- Quasi – Monte Carlo
- Stochastic Collocation
-
- etc.

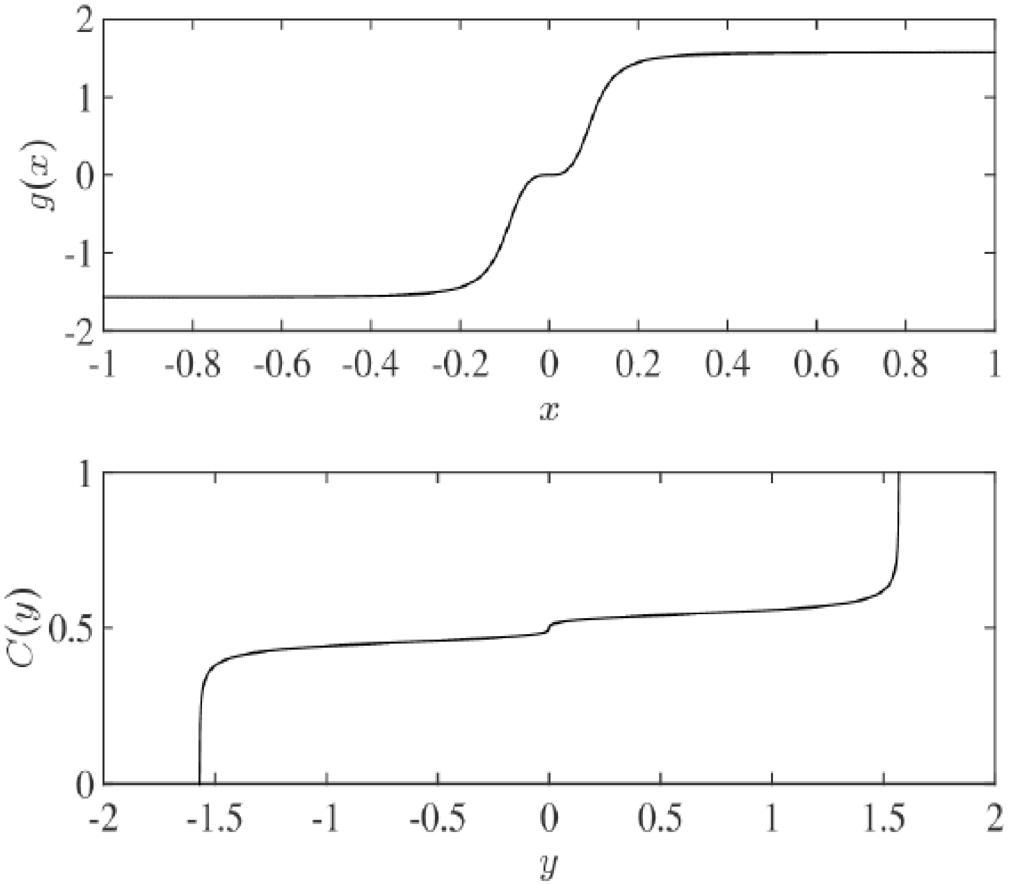
A simple adaptive method for sampling

Interpolant: radial basis function (mesh free)

Simple consideration:

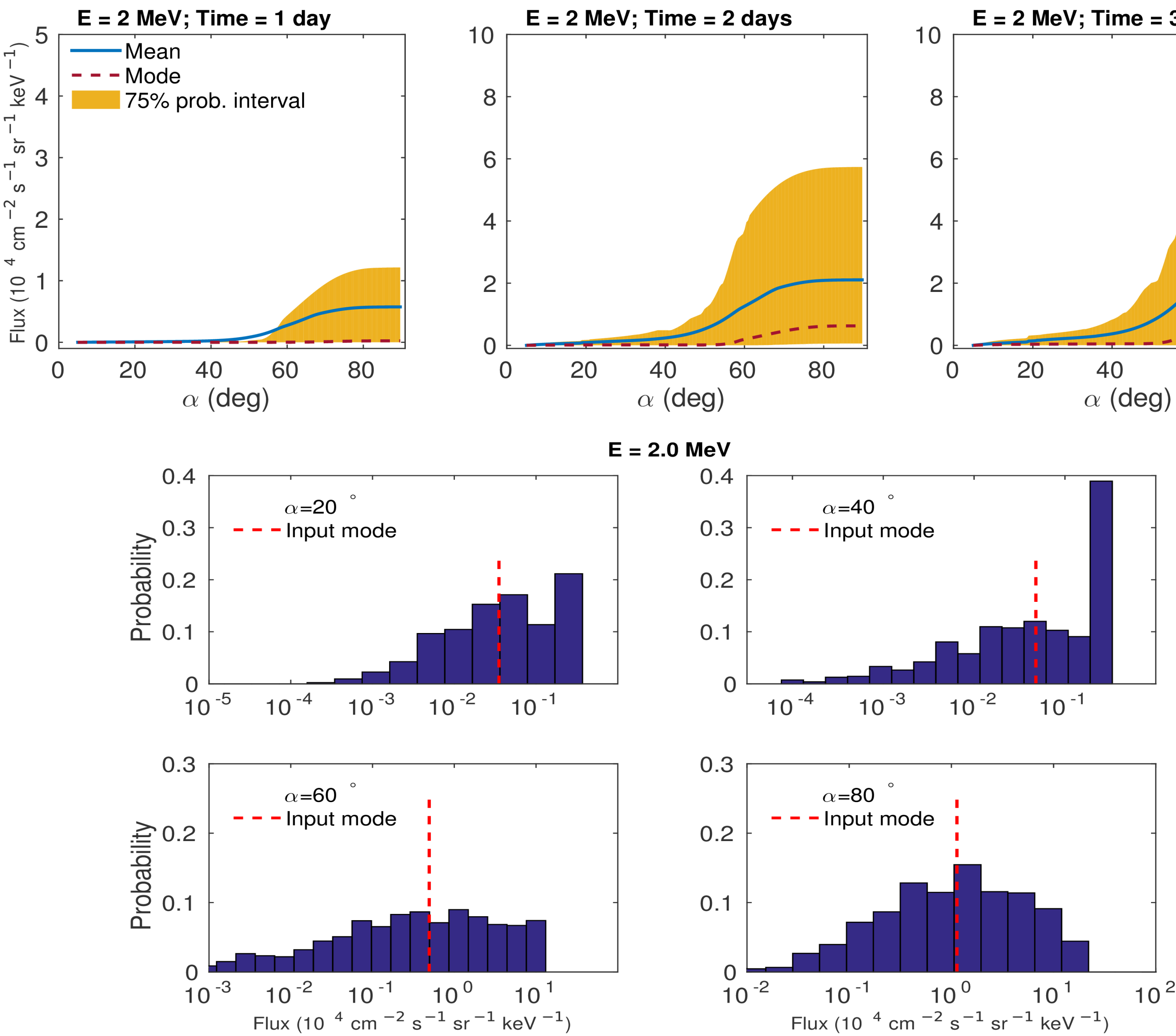
Large and small derivative of mapping function determine “flat” and “steep” regions in the cumulative distribution function.

This is “more or less” all you need to determine your sampling points!



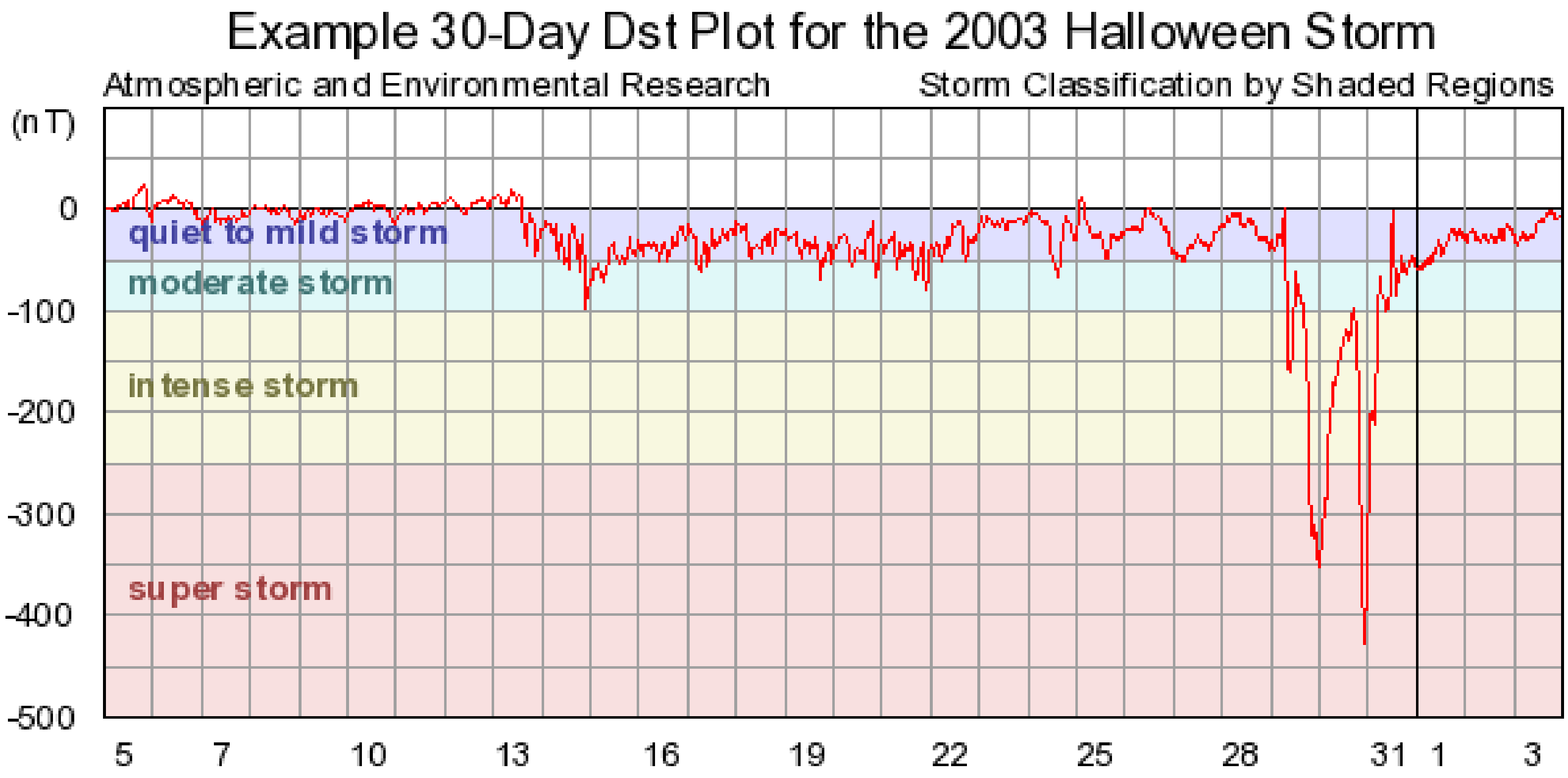
$g(x) = \arctan(10^3 x^3)$. *Top panel:* $g(x)$; *bottom panel: cdf* $C(y)$.

UQ on radiation belt simulations



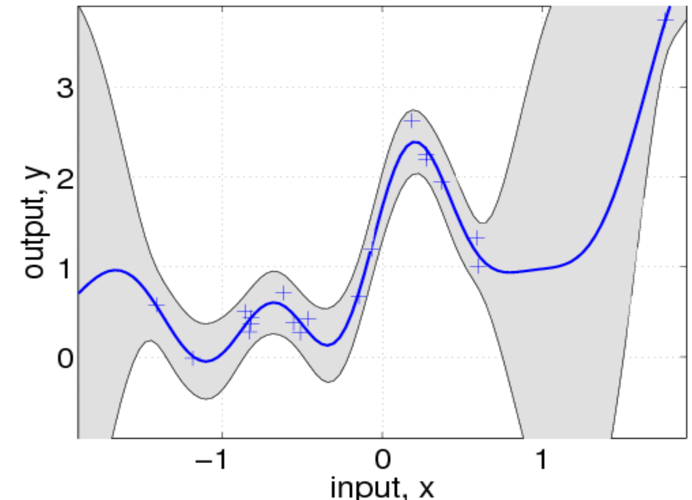
Gaussian Processes for forecasting the DST index

The DST index is a proxy for the geomagnetic activity



Gaussian Processes in a nutshell

We want to infer the values of an unknown function *f*. We assume that the conditional distribution of *f* with respect to input data is normally distributed.



$$\mathbf{f} = \begin{pmatrix} f(\mathbf{x}_1) \\ f(\mathbf{x}_2) \\ \vdots \\ f(\mathbf{x}_N) \end{pmatrix}$$

$$\mathbf{f}|\mathbf{x}_1, \cdots, \mathbf{x}_N \sim \mathcal{N}(\mu, \mathbf{\Lambda})$$

$$p(\mathbf{f} \mid \mathbf{x}_1, \cdots, \mathbf{x}_N) = \frac{1}{(2\pi)^{n/2} det(\mathbf{\Lambda})^{1/2}} exp\left(-\frac{1}{2}(\mathbf{f} - \mu)^T \mathbf{\Lambda}^{-1}(\mathbf{f} - \mu)\right)$$

$$\mu_i = \mathbb{E}[f(\mathbf{x}_i)] := m(\mathbf{x}_i)$$

$$\Lambda_{ij} = \mathbb{E}[(f(\mathbf{x}_i) - \mu_i)(f(\mathbf{x}_j) - \mu_j)] := K(\mathbf{x}_i, \mathbf{x}_j)$$

We need to specify the covariance matrix (kernel)

Gaussian Processes in a nutshell

The conditional distribution on new test points can be computed analytically (because we are dealing with Gaussian)

$$\mathbf{f}_*|\mathbf{X}, \mathbf{y}, \mathbf{X}_* \sim \mathcal{N}(\bar{\mathbf{f}}_*, \Sigma_*),$$

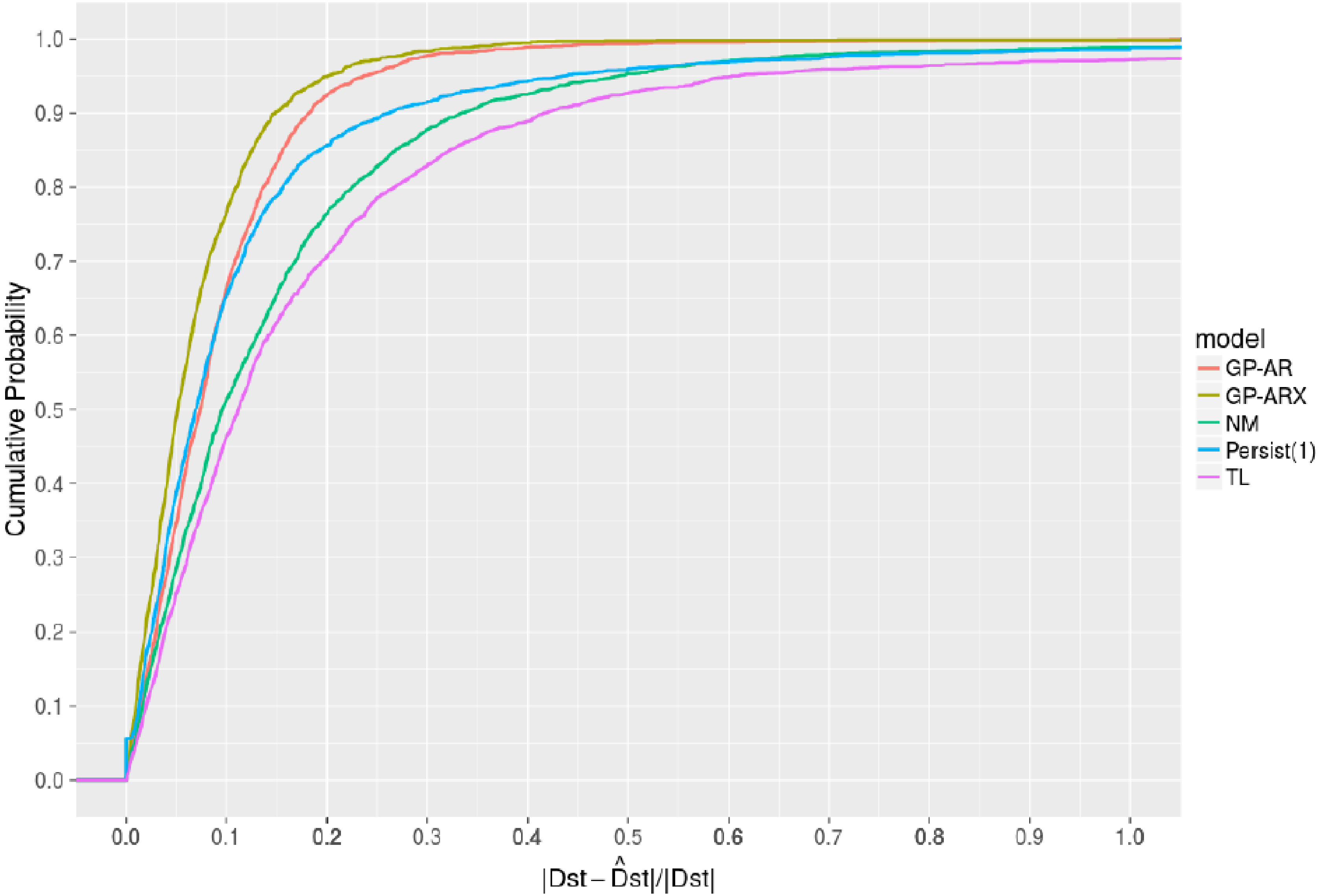
where

$$\bar{\mathbf{f}}_* = \mathbf{K}_*^T[\mathbf{K} + \sigma^2\mathbf{I}]^{-1}\mathbf{y}$$

$$\Sigma_* = \mathbf{K}_{**} - \mathbf{K}_*^T\left(\mathbf{K} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{K}_*$$

GP models do not make point predictions but output predictive (normal) distributions

CWI GP-ARX is the best model on the market!



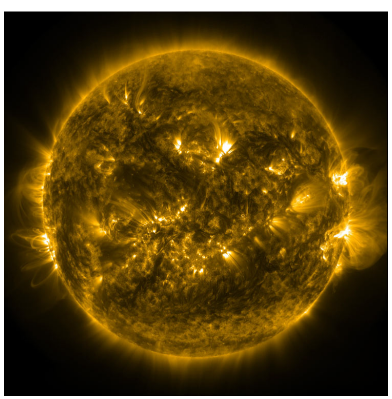
Future challenges

- Bayesian parameter estimation (inverse problem) to feed physics-based models
- Classification and regression of solar wind and magnetospheric conditions based on solar images

Future challenges

- Bayesian parameter estimation (inverse problem) to feed physics-based models
- Classification and regression of solar wind and magnetospheric conditions based on solar images

Dog!



CWI

80% chance of a geomagnetic storm!