# Classification in space science

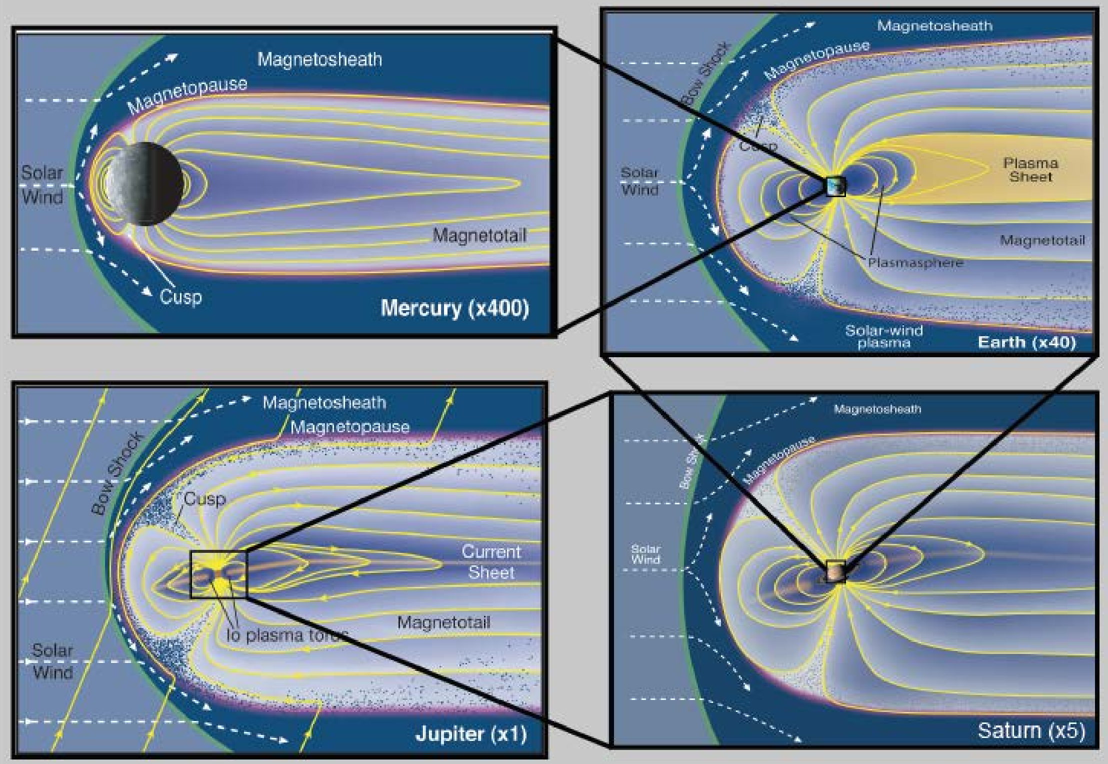
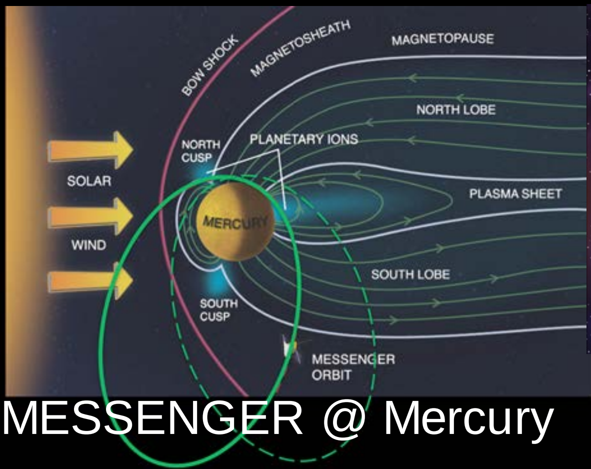



Image Courtesy Fran Bagenal and Steve Bartlett

Magnetic reconnection
- Products
- Textbook examples
- Steps to automation
- Current algorithms
- Outstanding problems
- Proposed architecture
- Cassini "challenge"!

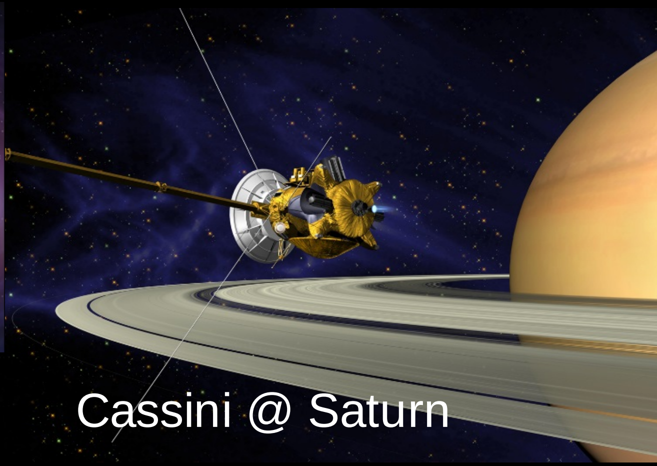

Dr. Caitríona Jackman c.jackman@soton.ac.uk
Manchester, November 2017
UNIVERSITY OF Southampton

# Motivation and Challenges:



MESSENGER @ Mercury


Juno @ Jupiter


Cassini @ Saturn

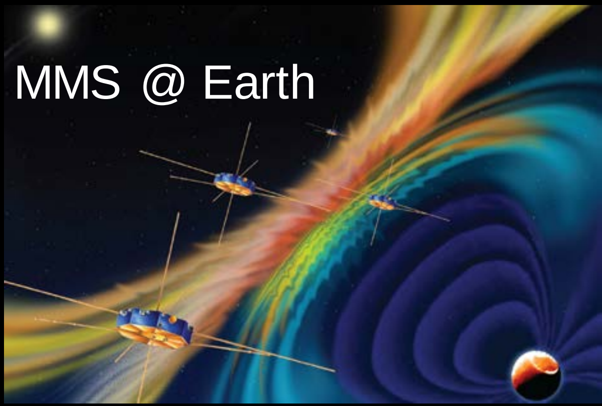
**Science gain from large statistical studies**

**Dataset availability**

**Mission funding windows**

**Bridge jargon gap**

**Time investment vs. reward**

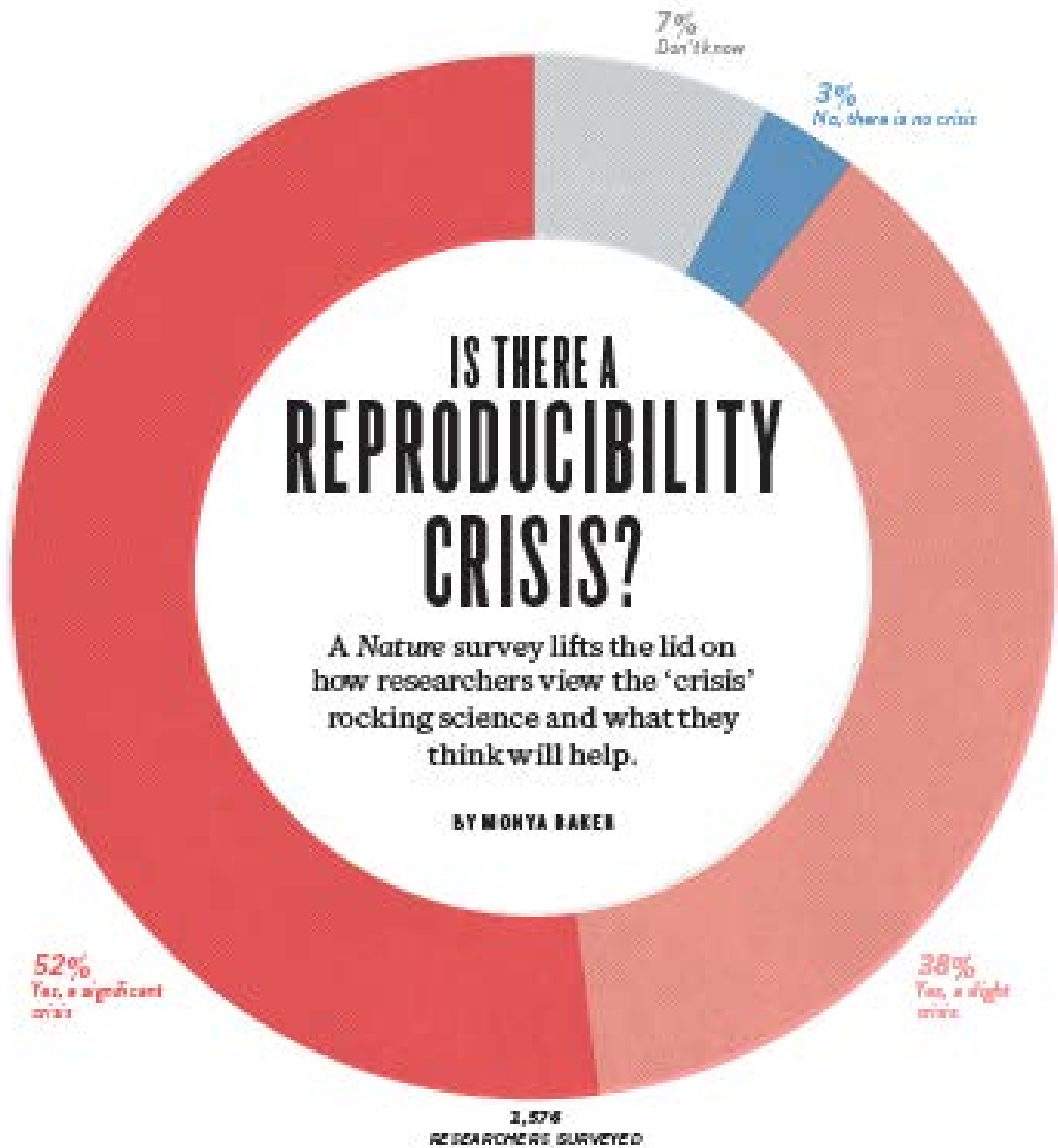

MMS @ Earth

Cluster @ Earth

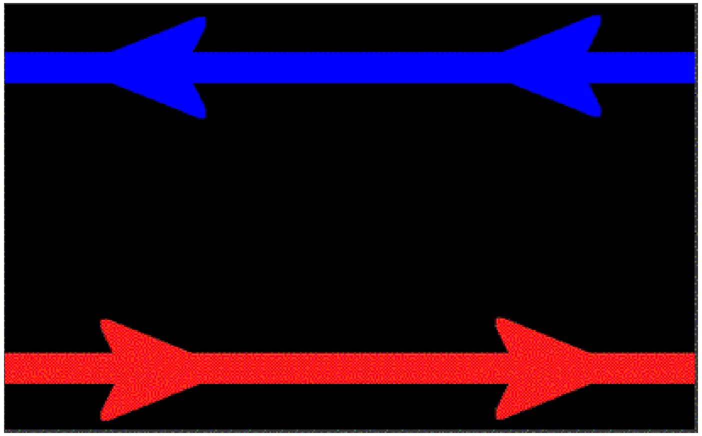
# Reproducibility:

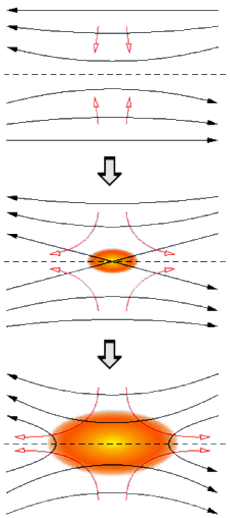
"More than 70% of researchers have tried and failed to reproduce another scientist's experiments…. And more than half have failed to reproduce their own experiments"

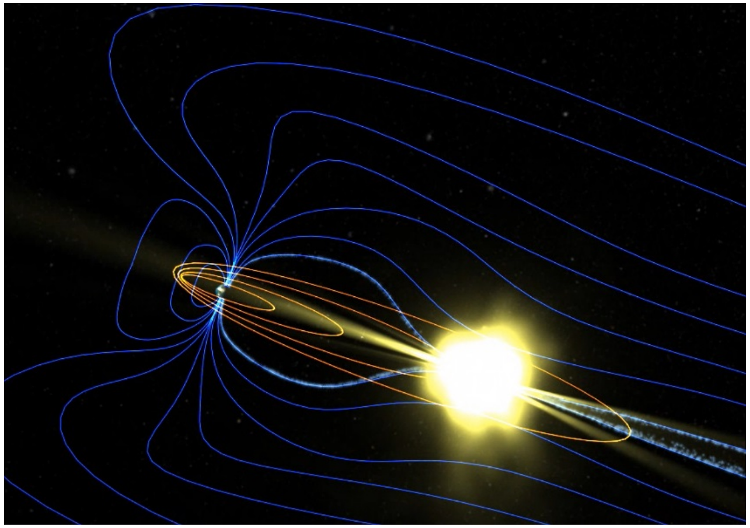
[Baker et al., Nature, 2016]

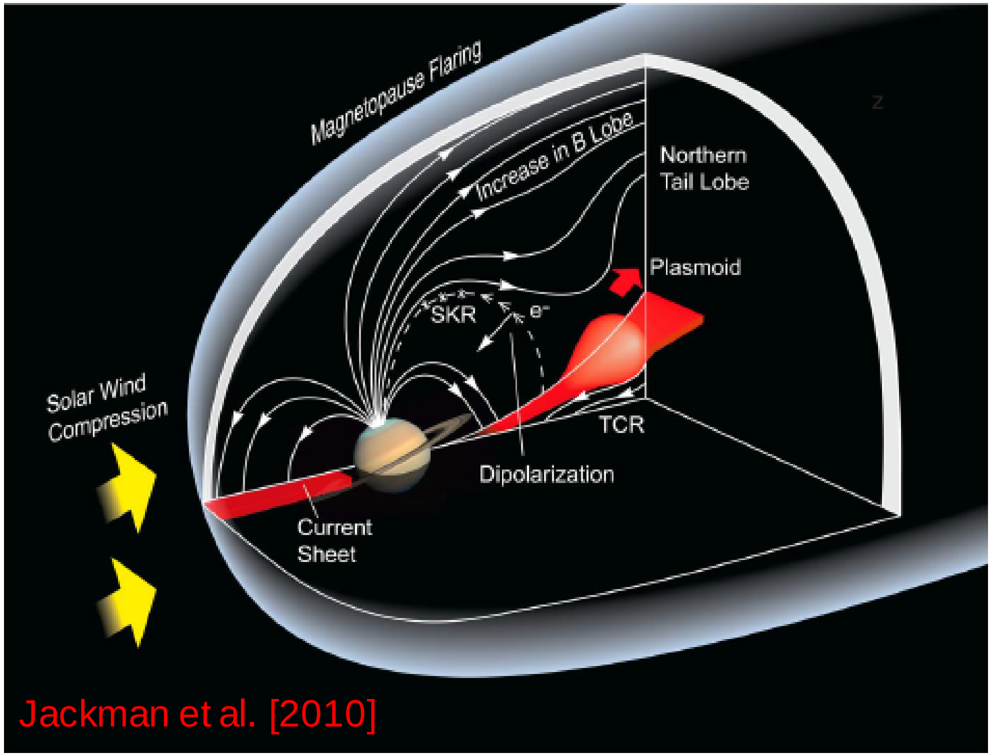


# Magnetic Reconnection









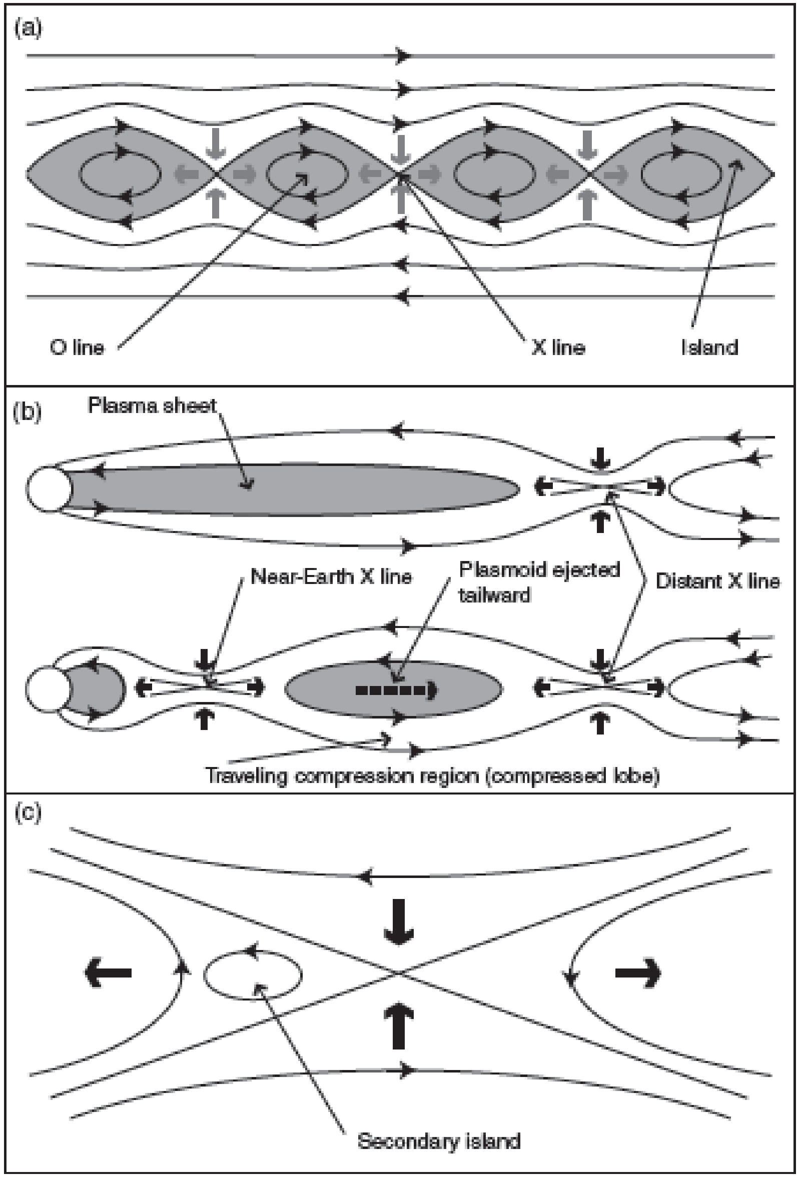

**Reconnection involves the explosive release of stored-up energy**

**Local:**

- Reconfiguration of magnetic field: plasmoid release/dipolarization
- Heating of plasma
- Change of plasma flow direction

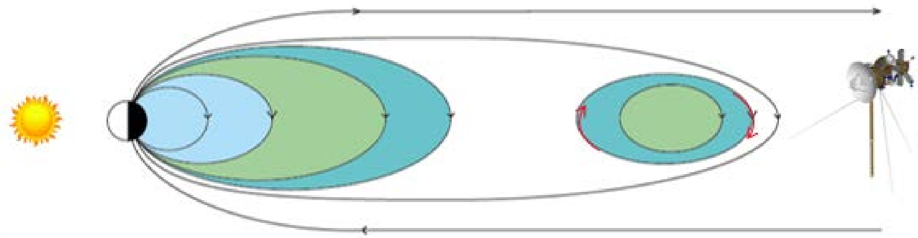
# In situ signatures of magnetic reconnection

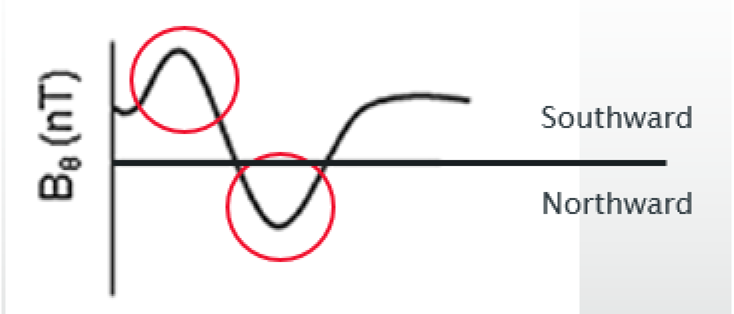
**EARTH**




Eastwood and Kiehas [2015]

**JUPITER/SATURN**

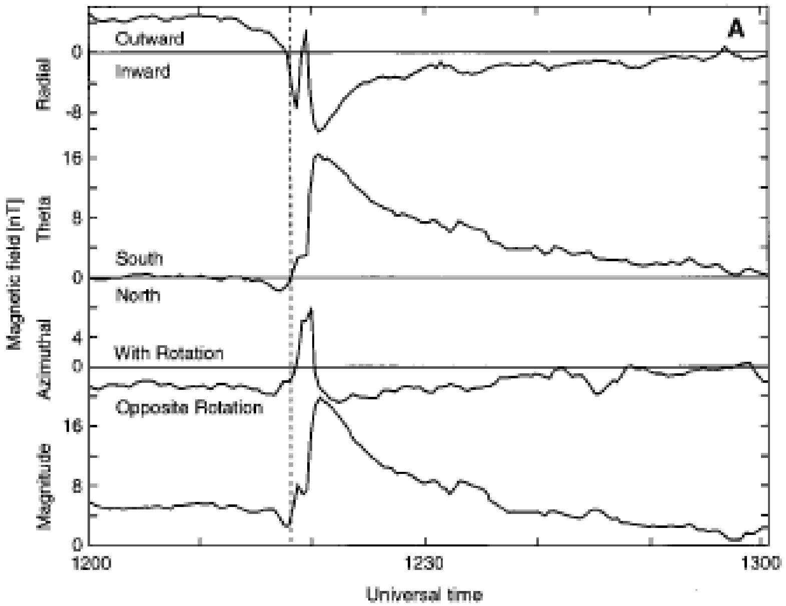




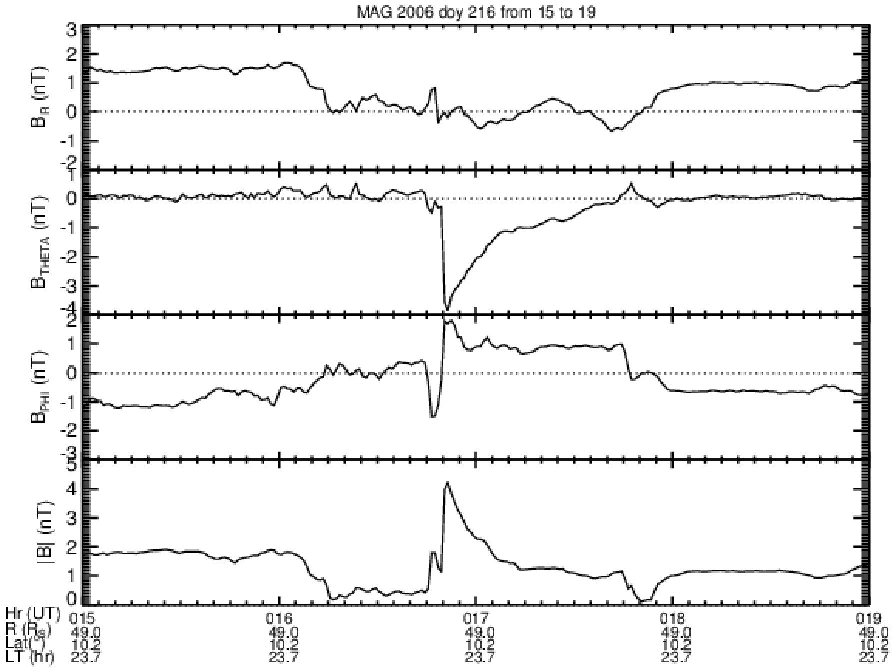

South-north turning of the field can indicate “plasmoid” tailward of the reconnection site

North-south turning of the field can indicate “dipolarization” planetward of the reconnection site

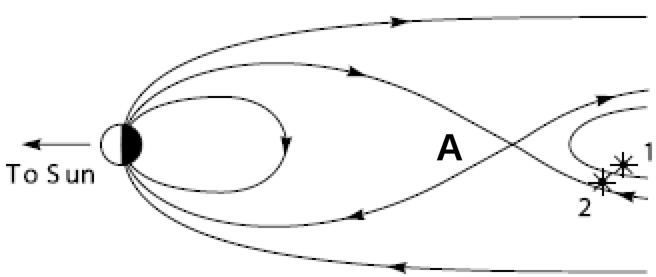
# Textbook examples of reconnection products




Jupiter dipolarization:
Russell et al. [Science, 1998]

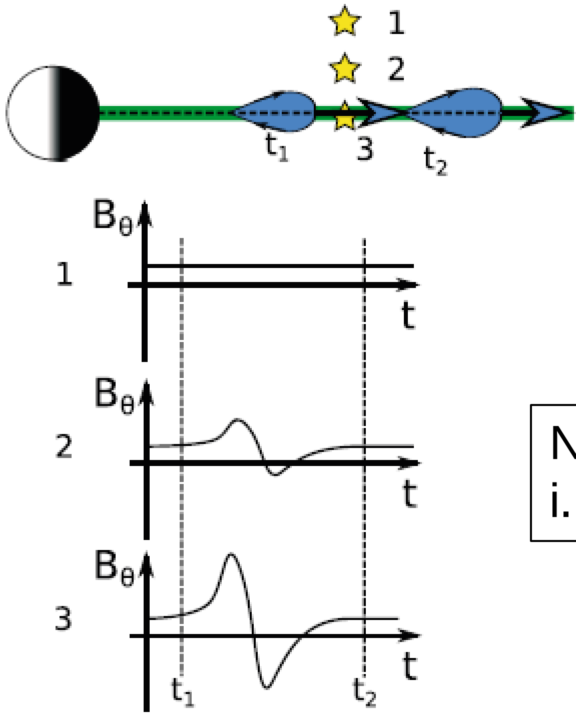



Saturn plasmoid: after
Jackman et al. [GRL, 2007]



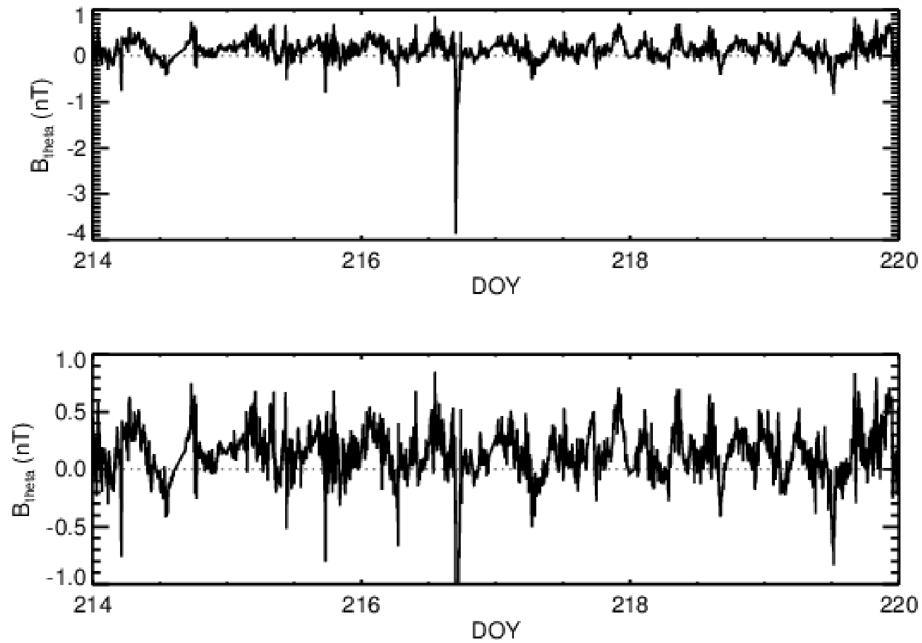

Early observations: “By-eye” selection of large south-north change in $B_\theta$.

# Challenge: Sensitivity of signature to viewing geometry




Not i.i.d?

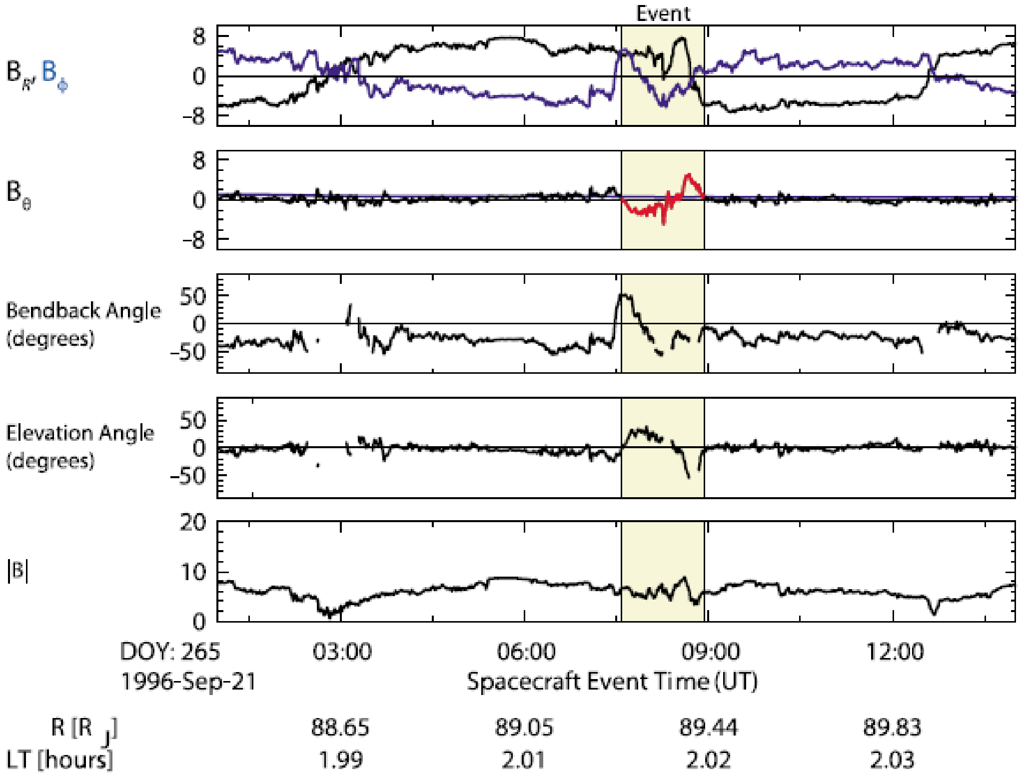
Smith et al. [2016]

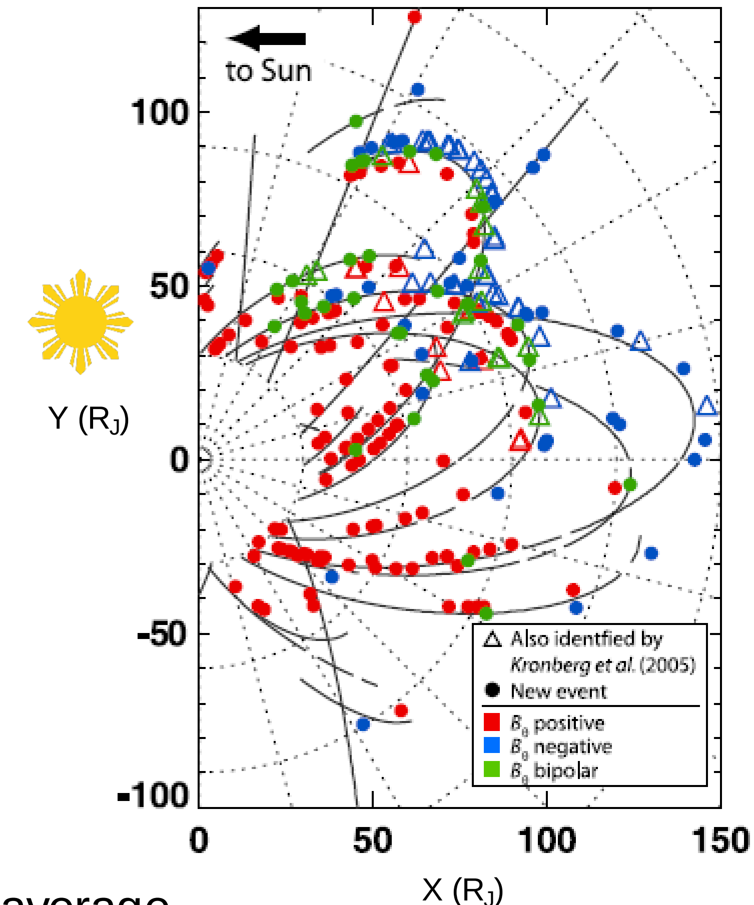



- Spacecraft may observe plasmoid passage from different latitudes
- Different penetration depth into structure
- Same sense but different magnitude of field change

- Plasmoids may also have different sizes
- Early “by-eye” studies found largest examples
- Many other smaller deflections also valid… noise vs. real features?

# Early attempts at automated searches: Jupiter




$$\frac{|B_\theta|}{\langle |B_\theta| \rangle} \geq 2,$$

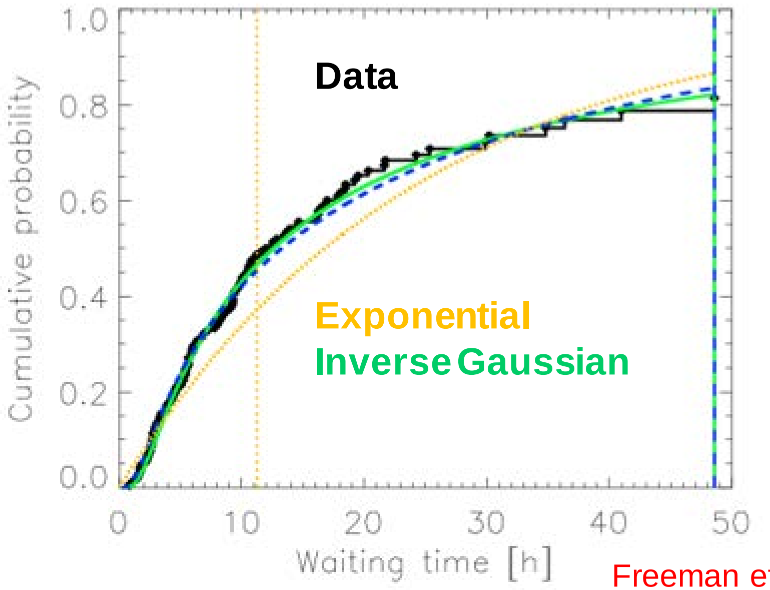
$B_\theta$ enhanced above 1-day running average

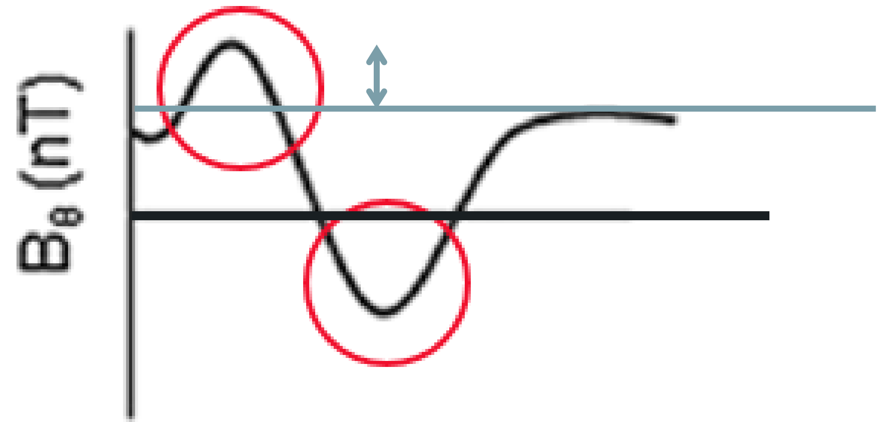
$B_\theta$ enhanced for at least 60 seconds

Vogt et al. [JGR, 2010]

Even basic automation gave an event catalogue of 249 events – ability to conduct statistical analysis

# Waiting time distribution of reconnection: Jupiter



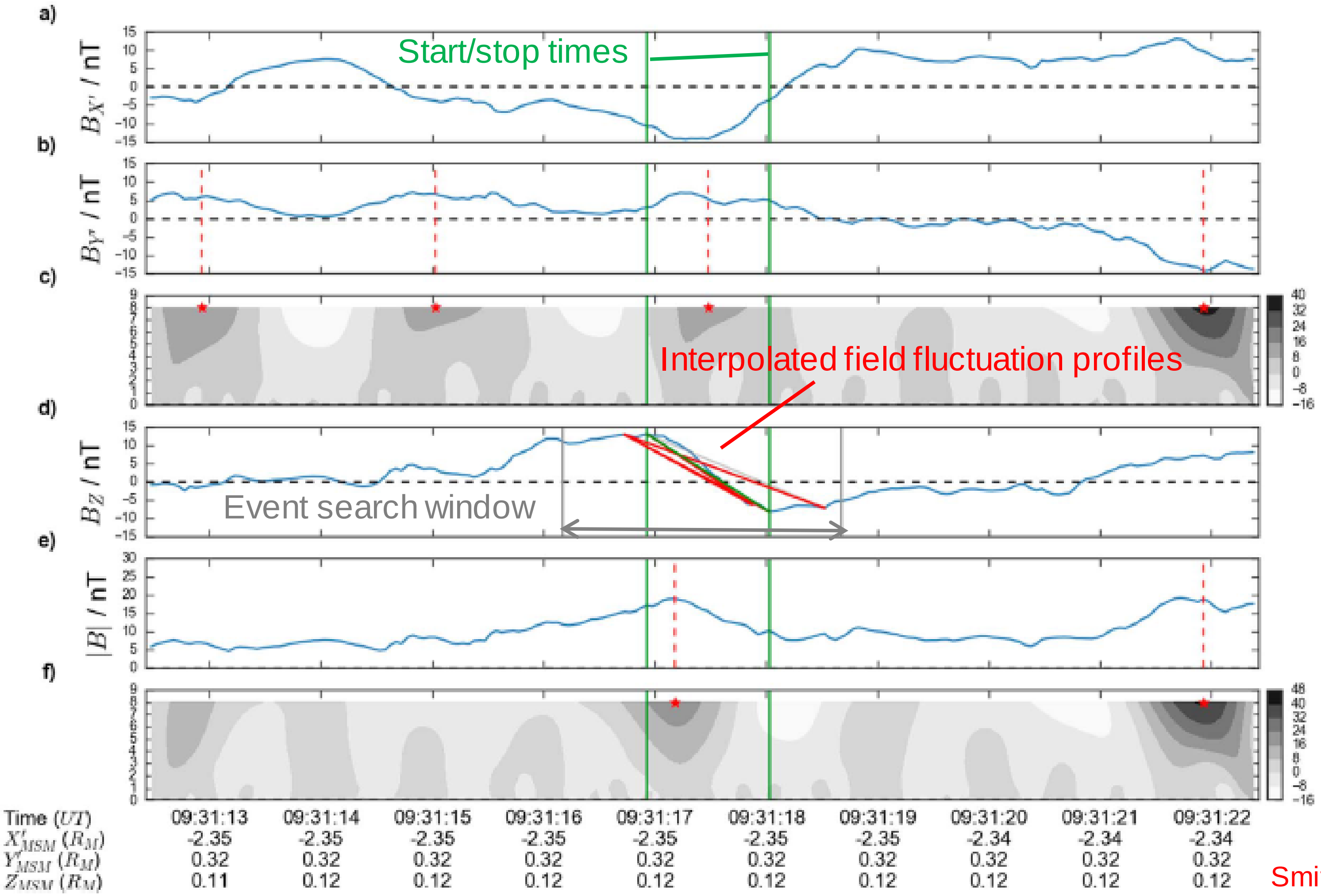

Freeman et al. [in prep., 2017]

Reconnection event waiting time distribution consistent with Inverse Gaussian… interpret as stochastic integrate-and-fire process.

Explore <u>sensitivity to reconnection event threshold</u>:
$|B_\theta|/ <|B_\theta|> = A$ $1.5 < A < 4.5$

When is an event "real" vs. "noise"? Can a ML algorithm help us to define this decision boundary?

# Increasing sophistication of automated searches: Mercury

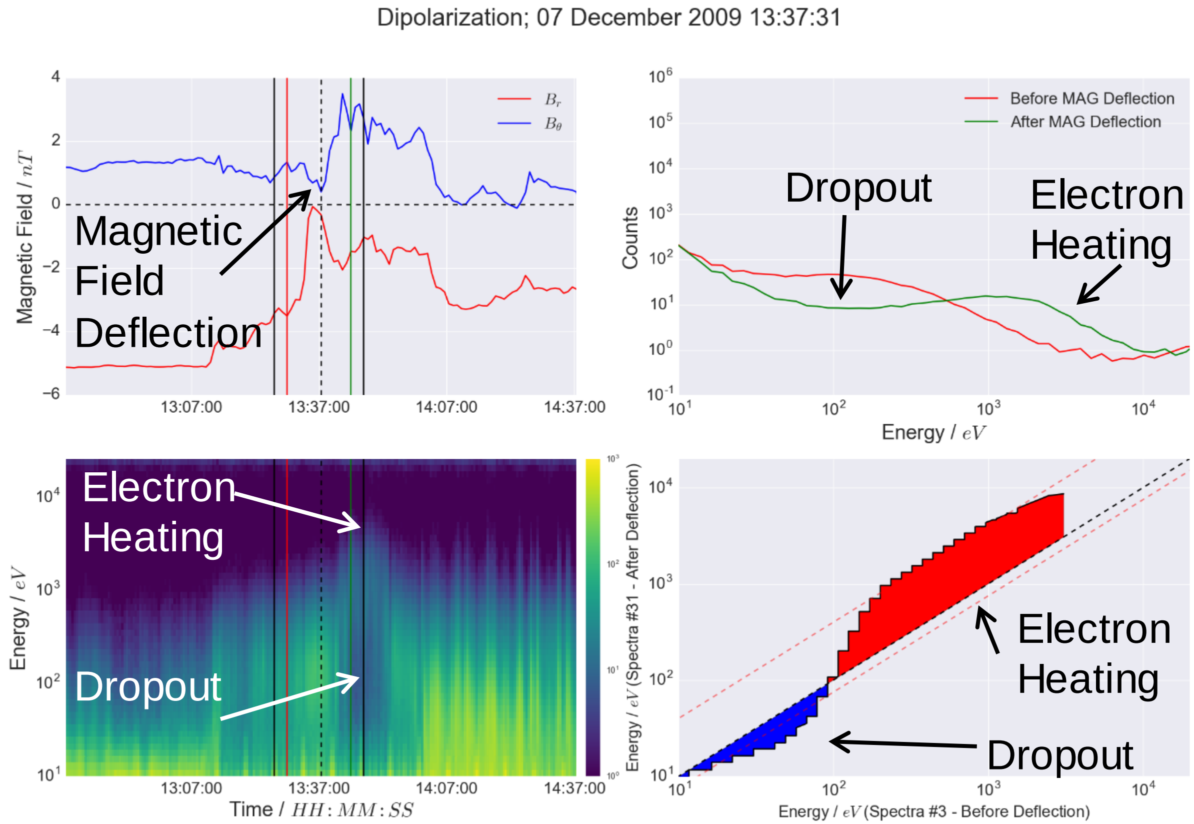



Smith et al. [2017a]

3-stage search for <u>force-free flux ropes</u>:

- Baseline crossing (threshold) & peak detection (continuous wavelet transform)
- Minimum Variance Analysis
- Fitting a force-free flux rope model

# Combined search of magnetic field and plasma data




Step 1:
Identify catalogue of magnetic deflections

Step 2:
Identify associated characteristic plasma signatures

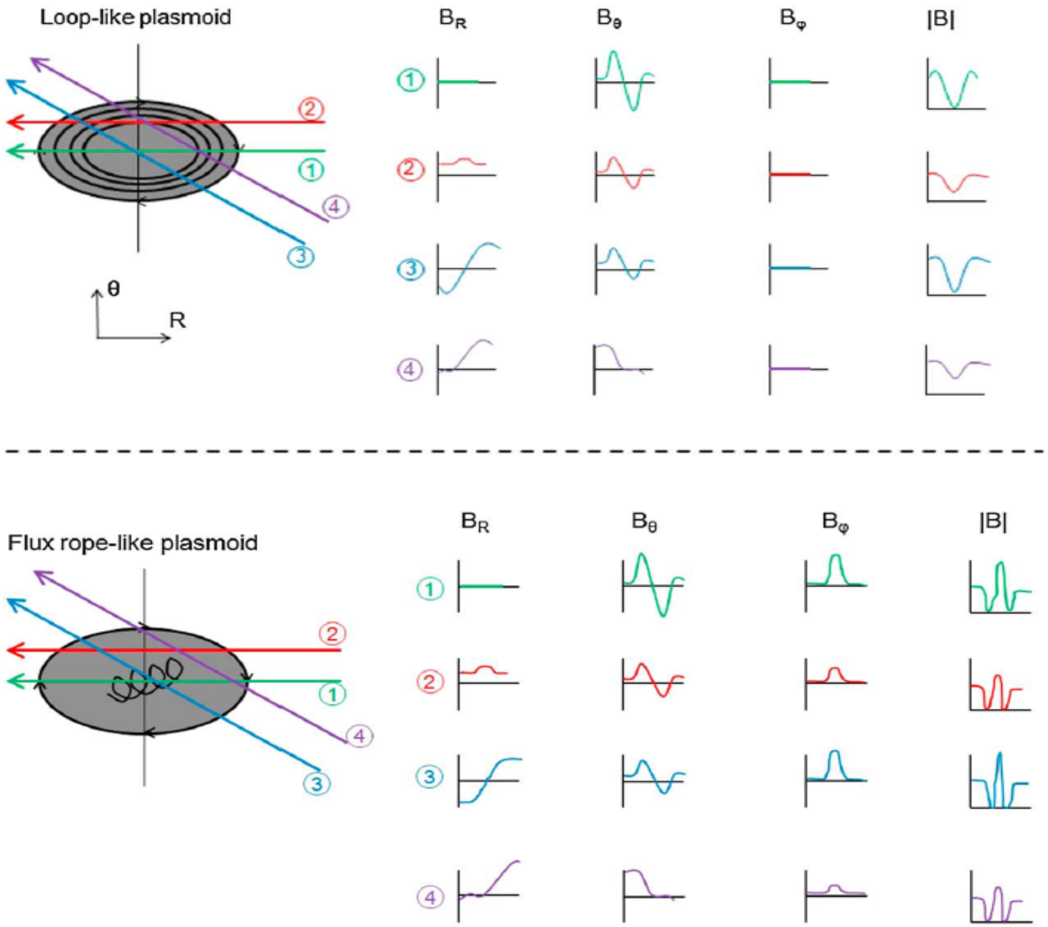
Use this to build the training set

Q-Q plotting: Gilchrist [2000]; Tindale and Chapman [2016]; Smith et al. [in prep. 2017b]

Automated search for electron heating and dropout:

- Define “initial”/background population
- Search sliding window after reconnection for “energized” population
- Maximise statistical difference using quantile-quantile plotting technique

# Further Challenges:




Jackman et al. [2014]

<u>"Zoo" of reconnection signatures</u> depending on:

- Interior structure of plasmoids/flux ropes
- Nature of spacecraft trajectory through the structure

Previous automated searches involve <u>"tuning"</u>:

- Selection of field deflection threshold
- Selection of time window to search over

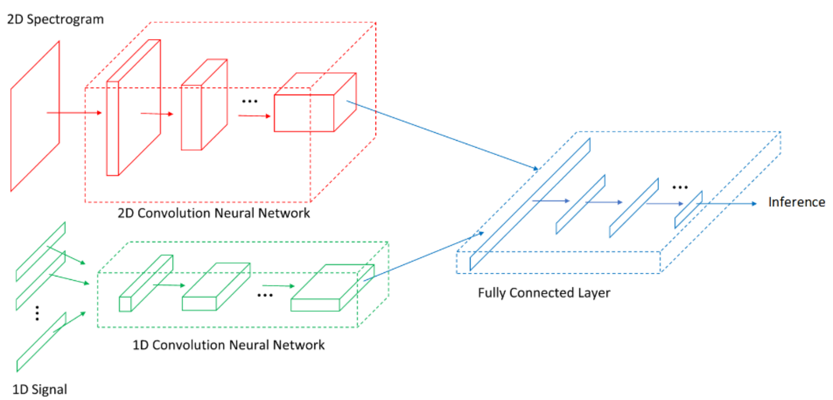
Aim to remove this manual intervention!

# Plan for automated classification scheme:

**Step 1: Data cleaning:**

(i) Ensure that the duration of the final data set and the cadence of measurements are appropriate to encompass the expected range of reconnection signature timescales

(ii) Ensure that gaps and anomalies are removed through filtering and interpolation




Jackman + Gunn, EPSRC proposal, 2017

**Step 2: Training set:**

(i) Build on previous algorithms and cross-check output (by eye?)

(ii) Divide data into training and test sets.

**Step 3: Construct algorithm**

(i) Convolutional neural network: combine 1-D magnetic field time series and the 2-D image-based plasma inputs.

(ii) Vary -D and 2-D kernels in the first layer over different scales, providing a mechanism for learning the key uncertainties in timescales.

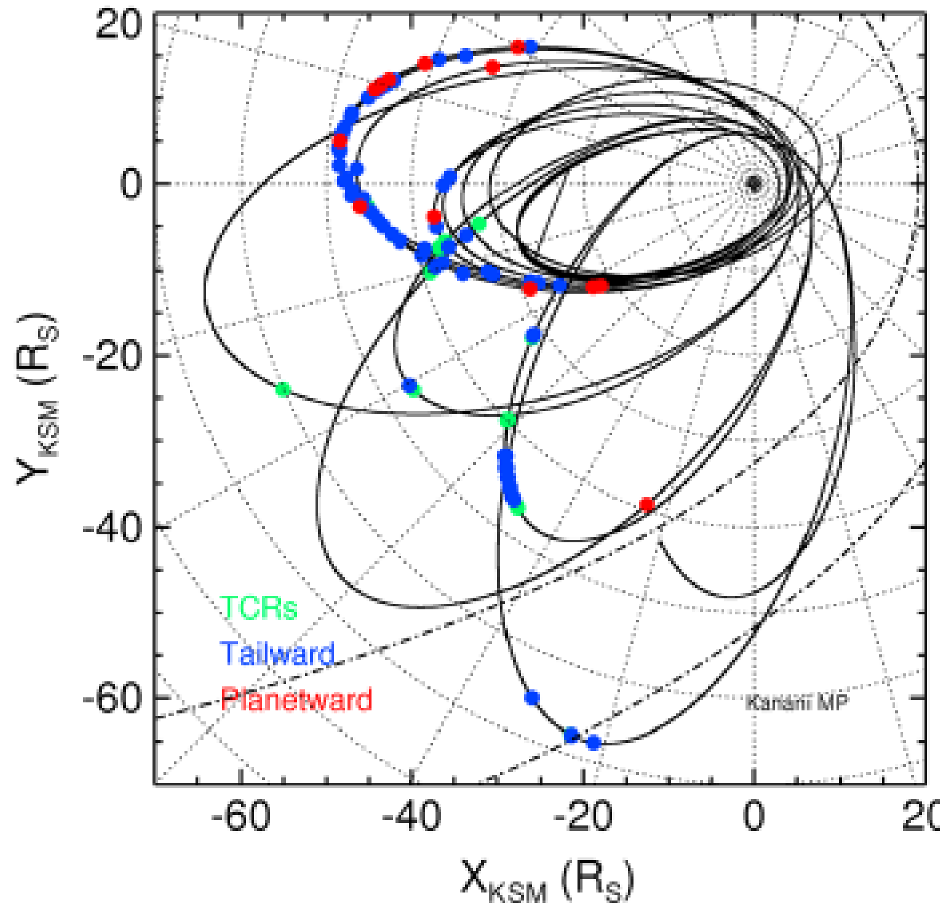
**Step 4: Evaluation**

(i) Class imbalance problem (sporadic reconnection)

(ii) Isolate 1-D and 2-D models to learn the extent to which a CNN can effectively combine 1-D and 2-D features

# Cassini data analysis challenge:

- 1 year of Cassini magnetometer data on Github
- 99 reconnection events labelled
- Can your ML algorithms find the same (or better) events?

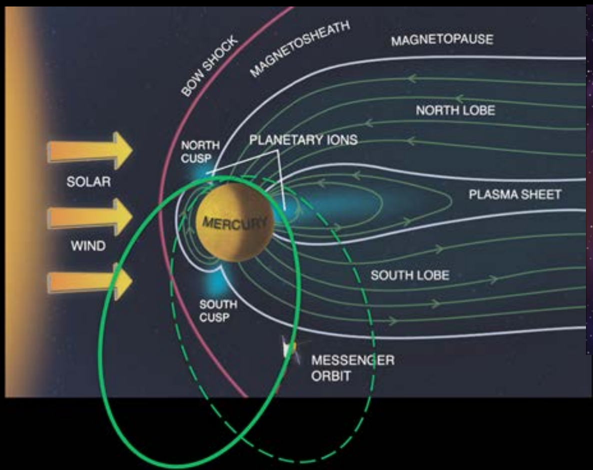


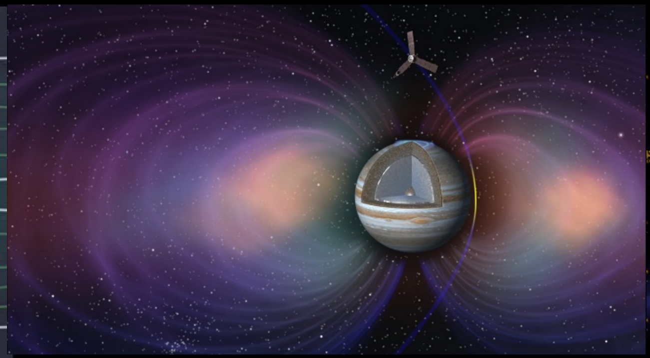

Jackman et al. [2014]

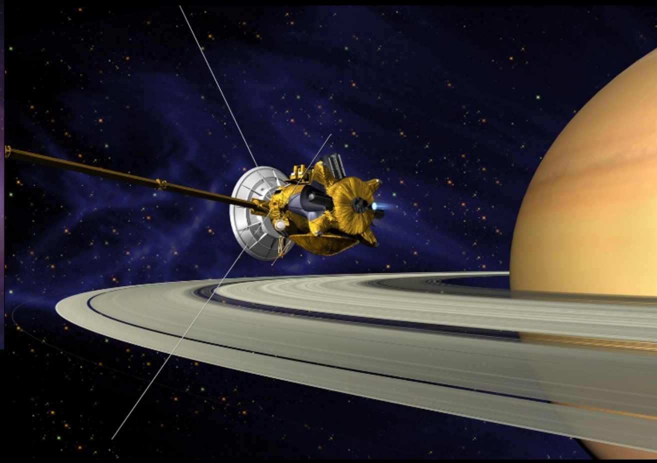
Email me:
c.jackman@soton.ac.uk

For details of data and labelling

# Summary:





**Need for ML algorithms to analyse vast amounts of magnetotail data**

**Binary selection. Viewing constraints and class imbalance problems to consider**

**Potential for large reward in terms of statistical understanding of influence of reconnection on magnetospheric dynamics.**