Information theory approaches in Sun-Earth's relations

Simon Wing The Johns Hopkins University



Outline

1. Information theory: mutual information, conditional mutual information, and transfer entropy

2. Three applications

2.1 Radiation belt dynamics



2.2 Solar dynamo



2.3 Radio waves at Saturn





Dependence between two variables



APL

Information theory



mutual information

Suppose that two variables x and y are binned so that they take on discrete values \hat{x} and \hat{y}

$$x \in \{\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n\} \equiv \aleph_1; \ y \in \{\hat{y}_1, \hat{y}_2, \cdots, \hat{y}_m\} \equiv \aleph_2$$

The variables may be thought of as letters in alphabets \aleph_1 and \aleph_2 , which have n and m letters

The entropy associated with each of the variables is defined as

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

where $p(\hat{x})$ is the probability of finding the word \hat{x} in the set of x-data and $p(\hat{y})$ is the probability

of finding word \mathcal{Y} in the set of y-data

entropy gives a measure of the amount of information in a variable/set (measure of disorder/uncertainty)

Shannon entropy

(Luis Serrano)



mutual information

The joint entropy is defined by

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

where $p(\hat{x}, \hat{y})$ is the probability of finding the word combination (\hat{x}, \hat{y}) in the set of (x, y) data

The mutual information: MI(x, y) = H(x) + H(y) - H(x, y)

MI is useful to identify nonlinear dependence between two variables [Tsonis, 2001]



MI(x,y) = information that x and y share

MI(x, y) = 0 iff x and y are independent

MI(x,y) = H(x) = H(y) if knowing x determines y

8/8/21



Mutual Information calculates that probability of knowing the y-value when the x-value is given.

Mutual Information Example

Heads	Tails	P _h	P _t	P _{hh}	P _{ht}	P _{th}	P _{tt}	MI
		1/2	1/2	10/46	12/46	12/46	12/46	.004
		1/2	1/2	22/46	1/46	1/46	22/46	0.82
		1/2	1/2	0	1/2	1/2	0	1.0
time	<u> </u>							

Same distributions, different mutual information!

Mutual information vs. Pearson and Spearman correlation





P = pearson correlation (linear relationship)

Sp = Spearman correlation (monotonic relationship)

MI = mutual information (linear and nonlinear relationship)

2(B/21

conditional mutual information

conditional mutual information (CMI) [Wyner, 1978] :

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

CMI(x, y | z) determines the mutual information between x and y given that z is known

if z is unrelated or random, CMI(x,y|z) = MI(x,y)

if x or y is known based on z, then CMI(x, y | z) = 0



11

An example of CMI: RB electron response to Vsw, nsw, and solar wind pdyn



12

transfer entropy

A common method to establish causal-relationships between two time series, e.g., $[x_t]$ and $[y_t]$, is to use a time-shifted correlation function

$$r(\tau) = \frac{\langle x_t | y_{t+\tau} \rangle - \langle x \rangle \langle y \rangle}{\sqrt{\langle x^2 \rangle - \langle x \rangle^2} \sqrt{\langle y^2 \rangle - \langle y \rangle^2}}$$

where r = correlation coefficient and τ = lag time

The results may not be clear if x and y have multiple peaks



transfer entropy

A better alternative is to use transfer entropy [Schreiber, 2000]

$$TE_{x \to y}(\tau) = \sum_{t} p(y_{t+\tau}, yp_t, x_t) \log\left(\frac{p(y_{t+\tau} \mid yp_t, x_t)}{p(y_{t+\tau} \mid yp_t)}\right)$$

 $yp_t = [y_t, y_{t-\Delta}, \dots, y_{t-k\Delta}], k+1 = \text{dimensionality of the system, and } \Delta = \text{first minimum in MI}$

 $TE(x \rightarrow y)$ gives a measure of how much additional information x provides in predicting the future of y beyond the degree to which y already predicts its own future

TE can be considered a special case of CMI

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

if no information flow from x to y, $TE(x \rightarrow y) = 0$ unlike correlation, $TE(x \rightarrow y) \neq TE(y \rightarrow x)$

Granger causality and transfer entropy

Granger causality principles (Granger, 1969; 1980):

- The cause occurs prior to its effect.
- The cause has *unique* information about the future values of its effect.

Relationship to transfer entropy

- X (Granger) causes Y if, in an appropriate statistical sense, X assists in predicting the future of Y beyond the degree to which Y already predicts its own future (Granger, 1969; Barrett et al., 2009)
- Transfer entropy reduces to Granger causality in a linear system
- Granger causality is pragmatic, well defined, and has delivered many insights into the functional connectivity of systems in a variety of fields (Ding et al., 2006, Set and Edelman, 2007, Cadotte et al., 2008).

Transfer entropy



Radiation belt dynamics

1. Radiation belt dynamics



2. Solar dynamo



3. Radio waves at Saturn



Magnetosphere



18

- radiation belts are highly variable
- to a large extent, solar wind controls the radiation belt variability
- it is not always clear which solar wind parameter plays the most dominant role





Space weather risks at geosynchronous orbit



• many geosynchronous satellites have terminated/died

Data set

- LANL geosynchronous satellite data 1989 2009
 - electrons with energy range 1.8–3.5 MeV
 - daily resolution: diurnal, MLT, lat-long dependences are reduced
 - Reeves et al. [2011] (<u>ftp://ftop.agu.org/apend/ja/2010ja0157535</u>)
- OMNI solar wind data at daily and hourly resolution
 - NASA OMNIweb (<u>http://omniweb.gsfc.nasa.gov</u>)
- Merged LANL and OMNI data set has 6438 data points

radiation belt electron flux vs. V_{sw}



 solar wind – radiation belt system is nonlinear (linear correlational analysis would be inadequate)

RB MeV J_e vs. V_{sw}



APL

RB MeV J_e vs. n_{sw}



APL

n_{sw} vs. V_{sw}





Input-Output problem



CMI can separate the effect of V_{sw} from n_{sw} and vice versa



CMI can separate the effect of V_{sw} from n_{sw} and vice versa



V_{sw} transfers ~2.7 times more information to RB electrons than n_{sw} does

CMI[Je(t + 0), $n_{sw}(t) | V_{sw}(t)] -$ CMI[Je(t + 0), sur[$n_{sw}(t)$] | $V_{sw}(t)$] = 0.09

"magnetopause shadowing" [Li et al., 2001, Shprits et al., 2006, Ukhorskiy et al., 2006] CMI[Je(t + 2), $V_{sw}(t) | n_{sw}(t)] -$ CMI[Je(t + 2), sur[$V_{sw}(t)$] | $n_{sw}(t)$] = 0.25

- acceleration caused by ULF waves generated at magnetopause by KHI [Reeves et al., 2007]
- "local acceleration" caused by VLF waves during particle injections [Summers et al., 1998; Horne et al., 2005]

Magnetosphere



31

The triangle distribution is well ordered by Transfer Entropy(TE)

RB MeV $J_e(t + 2 \text{ days}) \text{ vs. } V_{sw}(t)$



- Triangle distribution problem was first posed by Reeves et al. [2011]
- why is there such a huge variability when solar wind velocity is low?



High RB Je corresponds to high TE and vice versa

APL

RB Je lags both Vsw and nsw by 2 days

But, from the CMI analysis, RB Je response to nsw should have **0 day** lag

solar wind density effect on the triangle distribution



density gradient in the X direction because n_{sw} anticorrelates with V_{sw}

density gradient in the X and Y directions

high n_{sw} (or Pdyn) \rightarrow low RB J_e and vice versa



Ranking of solar wind parameters based on information transfer to RB J_e, given V_{sw}



solar wind and magnetospheric parameters ranked by information transfer to PSD given Vsw

rank	solar wind and magnetospheric parameters	peak information transfer (it _{max})	au (hours)
1	V _{sw}	0.13	46
2	SymH	0.035	30-70 (broad peak)
3	IMF [B]	0.033	6
4	P _{dyn}	0.024	7-10
5	AL	0.022	56
6	n _{sw}	0.020	11
7	IMF B _z < 0	0.018	4
8	IMF B _v	0.016	0
9	IMF $B_z > 0$	0.015	10
10	E _{sw}	0.014	46
11	σ (IMF B)	0.0051	4
Improving models with information theory

- 1. selecting input parameters based on information transfer
- 2. detecting changes in system dynamics with windowed TE
- 3. providing prediction horizon



Windowed TE can detect changes in system dynamics



windowed TE[$V_{sw}(t) \rightarrow J_e(t + 2 \text{ days})$]

use information theory and NN to model RB electron PSD

use information theory and NN to predict RB electrons

Neural Networks



Preliminary NN model results

input parameters: Vsw (t=0-48 hr), nsw(t=0-12 hr), IMF Bz(t=0-4 hr), IMF By(t=0-4 hr), AL(t=0-48 hr), Symh(t=0-48 hr), L*(t= 0 hr) output parameter: PSD(t=0) data: RBSP 2013-2018, μ = 700 MeV/G, I = 0.11 Re G^{0.5} work in progress!





The importance of SYM-H and AL in predicting RB PSD

without AL and SYM-H as input parameters



Summary

- 1. Information theory can help modeling by identifying
 - Relevant drivers of the radiation belt electrons
 - Response lag time
 - Prediction horizon (how far ahead can we predict?)
- 2. For ~1 MeV RB electrons (μ = 700 ± 25% MeV/G, I = 0.11 ± 25% Re G^{0.5}):
 - V_{sw} does not affect PSD at L* < 3.5
 - n_{sw} does not affect PSD at L* < 4.5
- 3. AL and Sym-H provide additional information to RB electrons beyond what Vsw provides
- 4. We rank solar wind and magnetospheric drivers of RB electrons based on CMI
- 5. We use NN to model RB electron PSD: PE = 0.66 (preliminary result)

Radiation belt dynamics

1. Radiation belt dynamics





3. Radio waves at Saturn







Introduction and motivation



- The roles of solar parameters in the the solar dynamo are not fully understood
- It is still a challenge to predict SSN

Babcock-Leighton type model [Babcock, 1961; Leighton, 1964; 1969]

Model predictions of SSN at solar max for solar cycle 24



It is still a challenge to predict sunspot number (SSN)

Data set

- SSN 1749–2016 SILSO website in Belgium
- sunspot area1874–2016 NASA MSFC website
- meridional flow 1986–2012 MWO [Ulrich, 2010] (from R. Ulrich)
- polar faculae 1906–2014 MWO, WSO, SOHO [Munoz-Jaramillo et al., 2012] at Solar Polar Fields Dataverse website
- polar field 1967–2015 MWO, SWO [Ulrich, 1992; Wang and Sheeley, 1995] (from Y.-M. Wang)
- axial dipole strength 1967–2015 MWO, WSO [Wang and Sheeley, 1995; 2009] (from Y.-M. Wang)
- aa index 1868–2010 NOAA NCEI website

All data are evaluated (averaged, interpolated) at monthly resolution

SSN and aa index

Hathaway et al. [1999]



Both, SSN and aa index exhibit cyclical variations

 $SSN \rightarrow$ aa index

aa index→ SSN [e.g., Ohl, 1966; Hathaway et al., 1999; Schatten and Pesnell, 1993; Wang and Sheeley, 2009]

Babcock-Leighton type model of solar dynamo



[Babcock, 1961; Leighton, 1964; 1969]

SSN and aa index



- peak $|corr(aa index(t), SSN(t + \tau))| \sim peak |corr(SSN(t), aa index(t + \tau))|$
- **But**, TE(SSN \rightarrow aa index) > TE(aa index \rightarrow SSN)
- more information is transferred from SSN to aa index than the other way around; aa index is a poor proxy for the solar polar field – this information cannot be obtained from correlational analysis

SSN and solar polar field



- TE(polar field \rightarrow SSN) peaks around $\tau \sim 30-40$ months, not 66 months assumed in some models
- peak significance = (peak TE mean(noise))/ σ (noise) = 19 σ
- TE(SSN \rightarrow polar field) is significant [Upton and Hathaway, 2014]

Introduction and motivation



Babcock-Leighton type model [Babcock, 1961; Leighton, 1964; 1969]

Which parameters control the polar field?



- surface flux transport models [e.g., Devore and Sheeley, 1987; Wang et al., 1989;2005] and flux transport dynamo models [e.g., Dikpati et al., 2006, Choudhuri et al., 2007]: meridional flow controls the strength of the polar field
- amount flux emergence (SSN) controls the polar field [Upton and Hathaway, 2014]
- TE(meridional flow \rightarrow polar field) peaks $\tau \sim 30-40$ (pos corr), $\sim 90-110$ months (neg corr)
- TE(SSN \rightarrow polar field) peaks $\tau \sim 50-80$ months (pos corr)

Are the past n cycles important for predicting SSN?



- Dikpati et al. [2004] suggested that meridional flow is slower at the bottom of the convection zone and hence the polar fields from the last 3 cycles should affect SSN (see also Charbonneau & Dikpati, 2000)
- TE(polar faculae → SSN) peaks at τ ~30-40 months but persists at a lower level thereafter for at least 400 months (~ 3 solar cycles)
- There are minima at $\tau \sim 1$ and ~ 2 solar cycle periods

Information transfer from polar field, polar faculae, and meridional flow to SSN



TE from polar faculae, polar field, and meridional flow to SSN 1986 - 2012

- noisy because data have shorter timespan, limited by meridional flow data
- TE([polar faculae, polar field]→SSN) > TE(meridional flow → SSN) at τ ~30–40 months, which may be consistent with Dikpati et al. [2010] model.
- TE(meridional flow → SSN) peaks around τ ~120 months (~1 solar cycle period), suggesting the meridional flow can be used to predict SSN one solar cycle period ahead



Conversion from toroidal to poloidal field is hard

Jie Jiang (Space Climate 7, 2019)



Comparing observations with Dikpati simulation



Summary

- $TE(SSN \rightarrow aa index) > TE(aa index \rightarrow SSN)$
- TE(polar field \rightarrow SSN) peaks at $\tau \sim 30$ -40 months (the response of SSN to polar field peaks ~ 3 -4 years, not 5.5 years).
- TE(polar faculae → SSN) peaks at τ ~30–40 months, but persists at lower level for at least 3 solar cycles
- Our results provide observational constraints to solar cycle models and theories

Radiation belt dynamics

1. Radiation belt dynamics

2. Solar dynamo

3. Radio waves at Saturn









RPWS radio wave data





1980-11-12 (317) 22:51:45 to 22:52:38

Cassini RPWS November 22, Day 324, 2003



Cassini INCA instrument

Cassini spacecraft

APL built the Ion and Neutral Camera (INCA) that imaged energetic neutral atoms (ENAs)



ENA Imaging



Cassini INCA ENA injections



Frame: SATURN 0.3010 Log cnts/quad (cm^{*}-s-keV)⁴ -0.3495-1.000

(UTC)

Seturn: SZS, SKR

Body shift 1919 secs linege shift 1919 secs Stare Ave: 16 Wth: 1 Rs 17.97 Lat -47.46 LT 0.02839.30L Lon____85.23

Particle injections and ENA emissions





If plasma and neutral sources are weak, radio emissions will not be periodic.



Type 1 and Type 2 ENA injections



- Type 1
 - > 10–12 Rs
 - wide MLT
 - intense
 - reconnection/current sheet collapse
- Type 2
 - < 10–12 Rs
 - narrower in MLT
 - less intense
 - interchange instability

Events

event	start time	end time	comment
1	2007 038 22:45:00	2007 039 17:25:00	
2	2007 042 02:00:00	2007 043 01:00:00	NB contaminated by SKR
3	2007 096 00:15:00	2007 096 17:22:00	
4	2008 025 14:30:00	2008 026 04:00:00	weak NB
5	2008 078 11:00:00	2008 079 21:00:00	weak NB
6	2009 012 12:45:00	2009 013 12:09:30	
7	2009 021 14:30:00	2009 023 13:45:00	
8	2009 065 06:00:00	2009 066 04:00:00	2 injections simultaneously
9	2009 112 22:20:00	2009 114 18:00:00	
10	2009 149 02:00:00	2009 149 16:00:00	
11	2009 151 00:00:00	2009 152 07:00:00	
12	2009 179 02:00:00	2009 179 17:51:00	
13	2009 181 00:00:00	2009 181 18:30:00	weak NB

- Search for type 2 injections that have RPWS wave data in ENA injection library 2007–2009
- 13 events found







RPWS radio wave data


2009 Apr 22 event



- $Corr(ENA(t), NB(t+\tau))$
- $Corr(ENA(t+\tau), NB(t))$

2009 Jan 21 event



Correlation vs. mutual information



APL

Particle injections and 5 kHz NB emissions

injection hot plasma \rightarrow temperature anisotropy \rightarrow upper hybrid waves \rightarrow density gradient at Enceladus density torus \rightarrow Z mode \rightarrow O mode









1. Magnetospheric Radio Emissions 1.6 Jupiter

- Hectometric Auroral Radiation (HOM)
 - Related to "energetic events" large-scale plasma heating events in the magnetotail (Louarn+2007, Woch+1998, Krupp+1998)
- Narrowband Emissions (n-KOM)
 - Source within Io Torus (Reiner+1993)
- JUICE will provide extensive observations in 2030





Simulated ENA image of magnetotail plasma heating at Jupiter. The ESA JUICE Mission to Jupiter will observe the system in ENAs and radio frequencies.

You are here!

brown dwarfs exoplanets

 $\sim 10^{22} - 10^{24}$ stars in the visible universe

Conclusion

- The phase relation between type 2 ENA injection and 5 kHZ narrowband emission is not random.
 - The 5 kHz narrowband emission lags ENA injection by a 60–120 min.
- Injections bring hot plasma to the inner magnetosphere which increase T anisotropy, which drives upper hybrid waves, which in the presence of steep density gradient at the outer edge of Enceladus density torus, mode convert to Z mode and then to O mode (NB is an O mode).
- The variability of the NB response lag time may be attributed to the relative position of Cassini with the injections
- Implications to NB emissions at Jupiter
- Implications to brown dwarfs and exoplanets

Summary

- Apply information theory to
 - Solar wind Radiation belt system
 - Solar dynamo
 - Radio waves at Saturn
- Information theory can be a useful tool for input-output problems
 - Establish linear and nonlinear correlations
 - Untangle the effects of input parameters that are correlated/anticorrelated with one another
- Information theory can be useful for modeling
 - Select and rank input parameters
 - Determine prediction horizons
 - Detect changes in the underlying dynamics of the systems