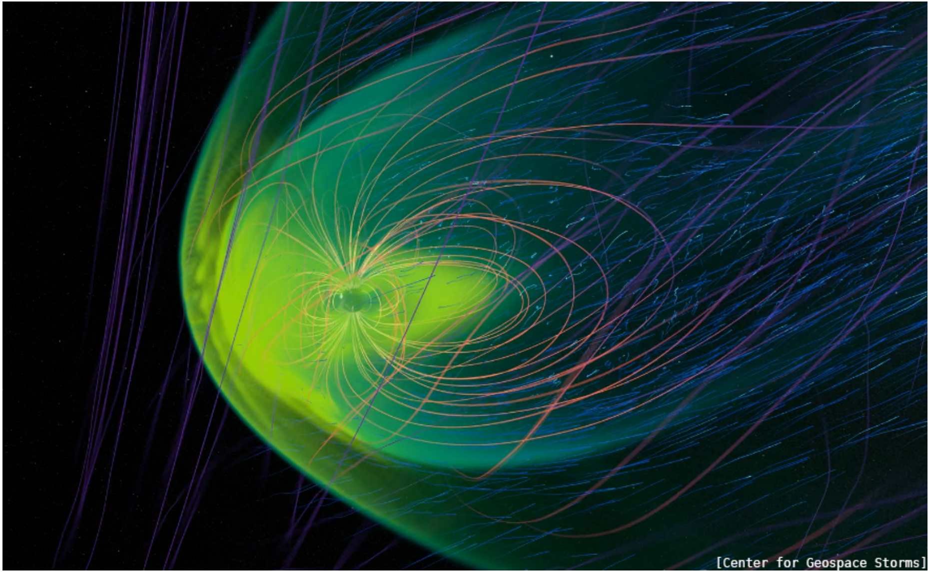
Correcting for measurement uncertainty in space weather

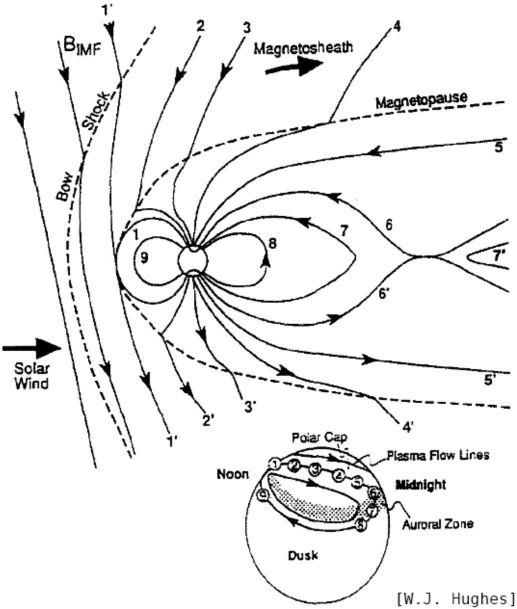
Magnetosheath control of the cross polar cap potential

Connor O'Brien[1], Brian Walsh[1], Evan Thomas[2]

November 8, 2024

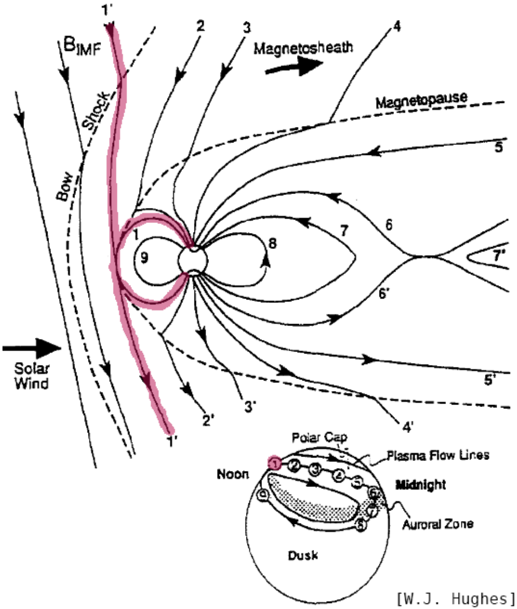
1: Center for Space Physics, Boston University, Boston, MA, USA
2: Thayer School of Engineering, Dartmouth College, Hanover, NH, USA


[Center for Geospace Storms]

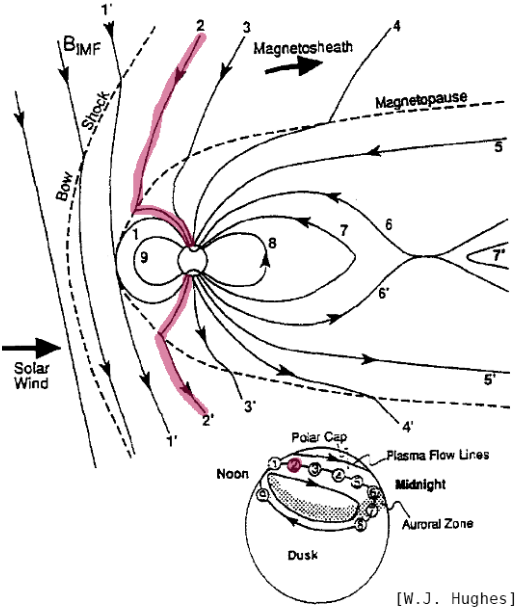


1'
B_{IMF}
2
3
4
Magnetosheath
Bow Shock
Magnetopause
5
6
7
8
9
1
7'
6'
Solar Wind
5'
1'
2'
3'
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

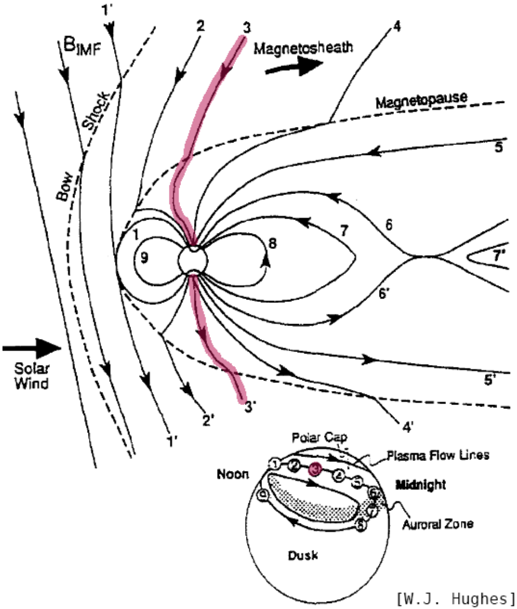
[W.J. Hughes]


1'
B_IMF
2
3
4
Magnetosheath
Shock
Bow
Magnetopause
5
6
7
8
1
9
7'
6'
Solar Wind
5'
3'
2'
1'
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

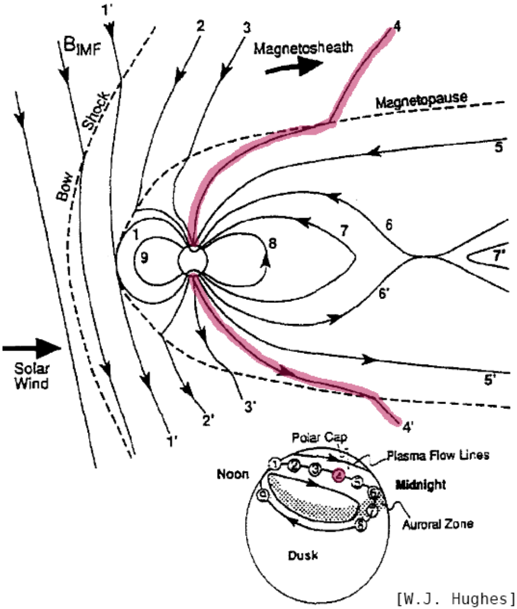
[W.J. Hughes]


1'
2
3
4
B_{IMF}
Magnetosheath
Magnetopause
Shock
Bow
5
6
7
8
1
9
7'
6'
5'
Solar
Wind
3'
2'
1'
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

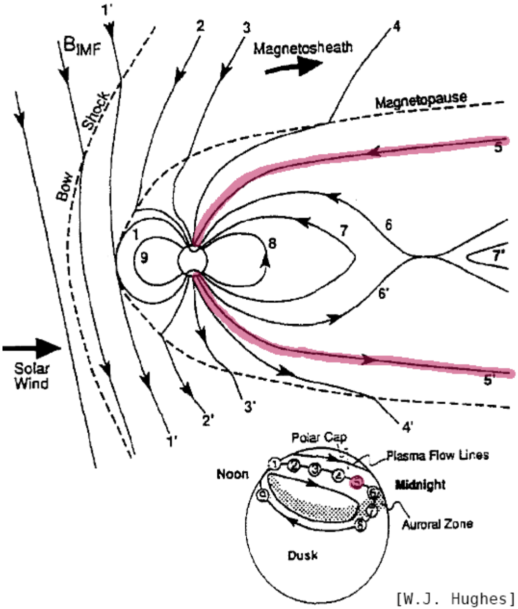
[W.J. Hughes]


1'
B_IMF
2
3
Magnetosheath
4
Shock
Magnetopause
5
Bow
1
9
8
7
6
7'
6'
Solar Wind
5'
2'
3'
1'
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

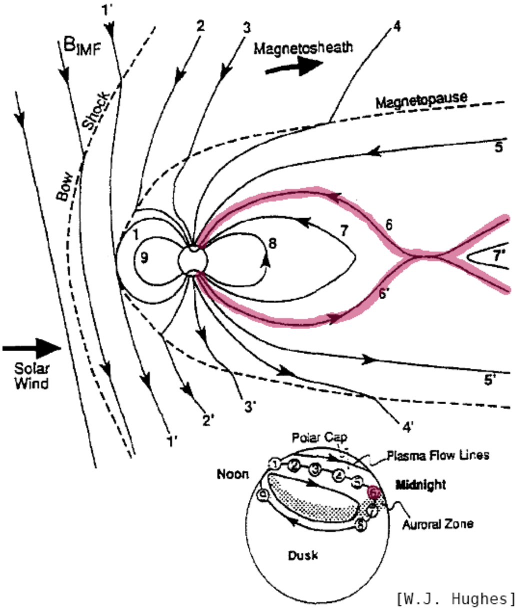
[W.J. Hughes]


1'
B_{IMF}
2
3
Magnetosheath
Shock
Magnetopause
Bow
5
6
7
8
1
9
7'
6'
Solar Wind
5'
2'
3'
1'
4
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

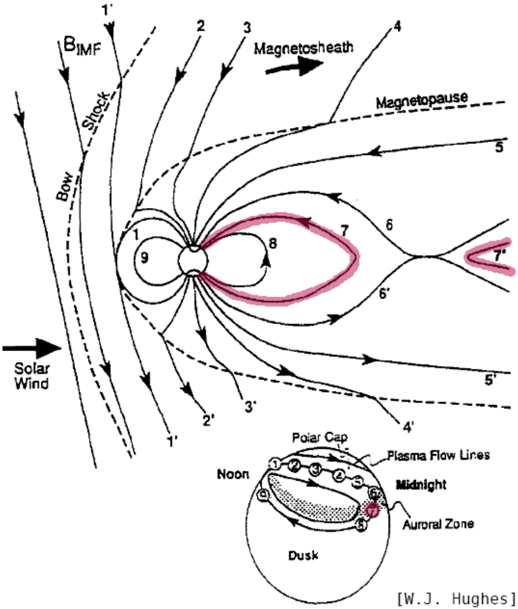
[W.J. Hughes]


1'
B_IMF
2
3
4
Magnetosheath
Magnetopause
Shock
Bow
5
6
7
8
1
9
7'
6'
5'
Solar Wind
2'
3'
4'
1'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

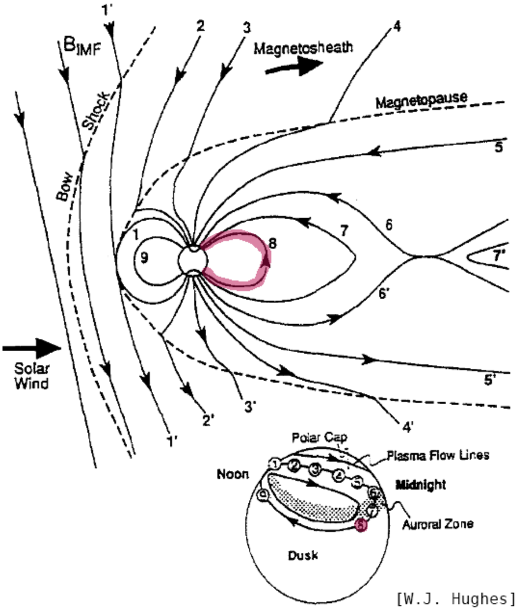
[W.J. Hughes]


1'
B_IMF
2
3
4
Magnetosheath
Magnetopause
Shock
Bow
5
6
7
8
1
9
7'
6'
Solar Wind
5'
1'
2'
3'
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

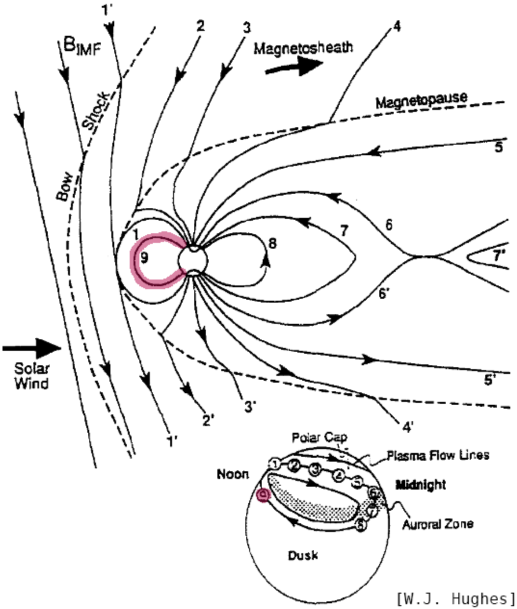
[W.J. Hughes]


1'
B_{IMF}
2
3
4
Magnetosheath
Magnetopause
Bow Shock
5
6
7
8
1
9
7'
6'
Solar Wind
5'
3'
2'
1'
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

[W.J. Hughes]

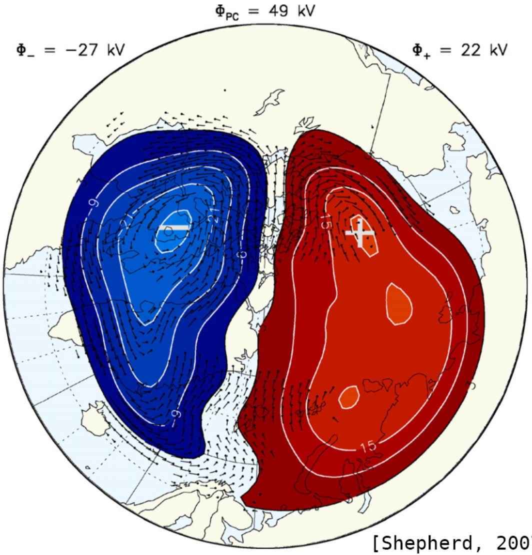

1'
B_IMF
2
3
Magnetosheath
4
Magnetopause
Shock
Bow
5
6
7
8
1
9
7'
6'
Solar Wind
5'
3'
2'
1'
4'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

[W.J. Hughes]

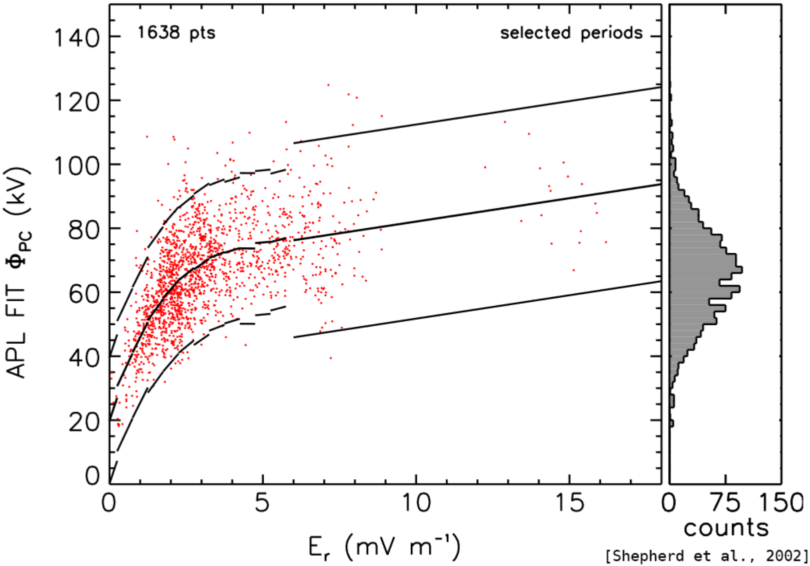

B_IMF
1'
2
3
4
Magnetosheath
Magnetopause
Shock
Bow
5
6
7
8
9
1
7'
6'
Solar Wind
5'
4'
3'
2'
1'
Polar Cap
Plasma Flow Lines
Noon
Midnight
Auroral Zone
Dusk

[W.J. Hughes]

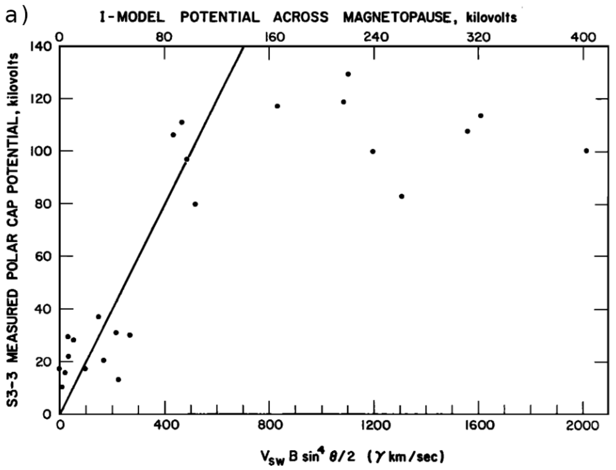
- Plasma moving perpendicular to $\vec{B}$ in a collisionless environment necessitates a convection electric field (Ohm's law): $\vec{E} = \vec{v} \times \vec{B}$.
- The size of the convection potential is called the polar cap potential ϕ_{PC}.
- This electric field is *notionally* the reconnection electric field at the magnetopause applied to the ionosphere.
- **How does ϕ_{PC} respond to the solar wind electric field?**

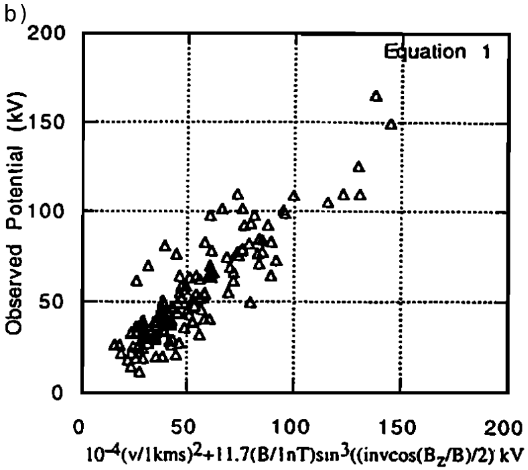


[Shepherd, 2007]



[Shepherd et al., 2002]





a) Wygant et al. 1983, saturation at $E_M = 0.5mV/m$ b) Boyle et al. 1997, no saturation

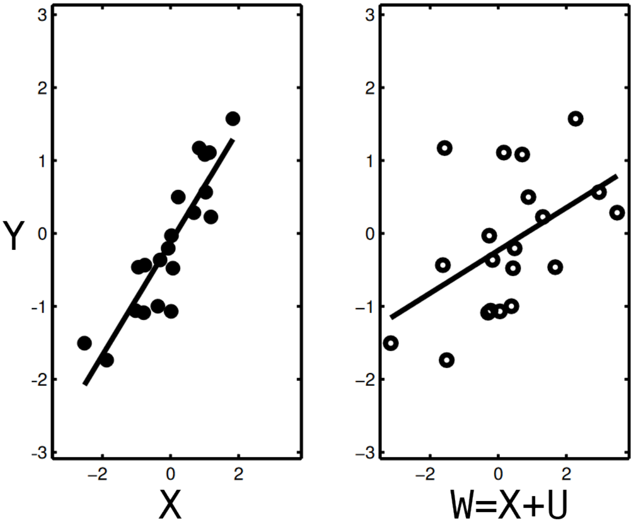
Consider the dependent variable Y

$$Y = \beta_X X + \beta_0 + \epsilon \tag{1}$$

Where ϵ is some random error.
When trying to determine β_X and β_0, we can't regress Y on X. We only have access to the error prone measurement of X we call W:

$$W = X + U \tag{2}$$

If U is uncorrelated with X, regressing Y on W gives an underestimate of β_X.

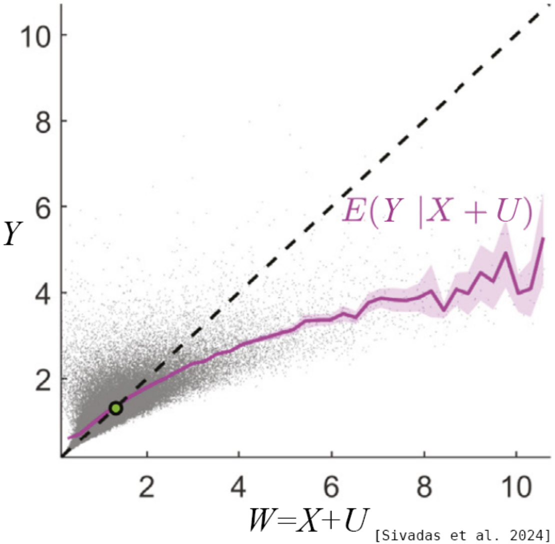


[Carroll, 2006]

If U is correlated with X, regressing Y on W creates an apparent nonlinear relationship between Y and W.

The correlation between X and U controls the amount and direction of the bias.

Could this bias from measurement uncertainty be responsible for disagreement about ϕ_{PC} saturation?



[Sivadas et al. 2024]

What part of the variation of ϕ_{PC} with respect to solar wind driving is physics, and what part is just uncertainty?

Solar Wind and Magnetosheath Data

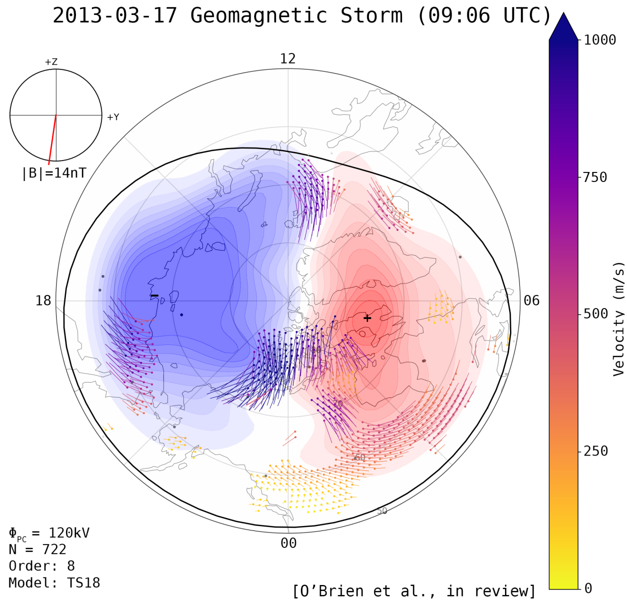
PRIME

- Probabilistic solar wind prediction algorithm
- Predicts solar wind just upstream of Earth's bow shock using solar wind data from L1
- Propagation location: $13.25R_E$ upstream of Earth on Sun-Earth line
- Used to obtain 2min-cadence $\vec{v}_{ion}$, $\vec{B}$, and n_E with uncertainties from 2013-2022

PRIME[SH]

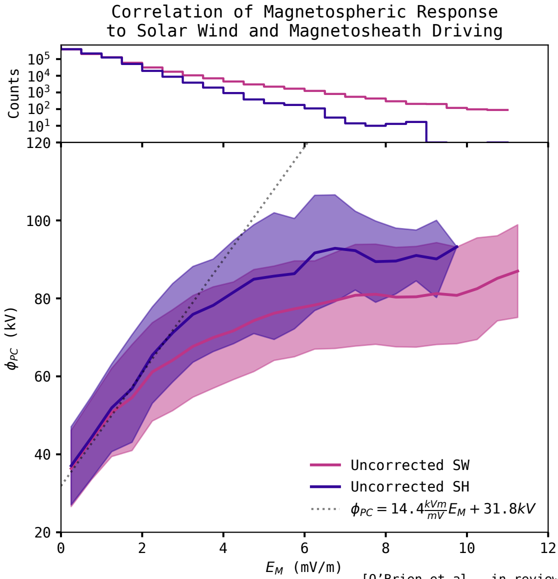
- Probabilistic magnetosheath prediction algorithm
- Predicts plasma and magnetic field in dayside magnetosheath using solar wind data from L1
- Propagation location: Middle of magnetosheath $0.5 - 1R_E$ downstream of bow shock
- Used to obtain 2min-cadence $\vec{v}_{ion}$, $\vec{B}$, n_{ion}, $T_\perp$, and $T_{||}$ with uncertainties from 2013-2022

From both datasets, we calculate $\vec{E}_M = v_X B_Z$ in GSM coordinates.

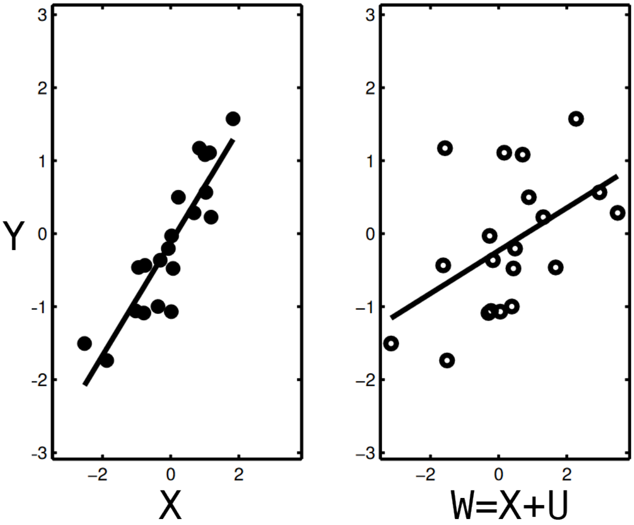


- SuperDARN is a network of 30+ HF coherent backscatter radars in the northern and southern hemispheres.
- SuperDARN is used to measure plasma convection vectors once every 2 minutes for 10 years (2013-2022).
- Then, spherical harmonics are fit to the convection maps to get ϕ_{PC}.

- ϕ_{PC} is correlated with the solar wind and magnetosheath E_M
- An initially linear relationship (dotted line) becomes nonlinear and saturates for both regimes
- ϕ_{PC} saturates with respect to solar wind $E_M > 3.5mV/m$
- ϕ_{PC} saturates with respect to magnetosheath $E_M > 4.5mV/m$
- **Is this a symptom of regression bias?**



[O'Brien et al., in review]



[Carroll, 2006]

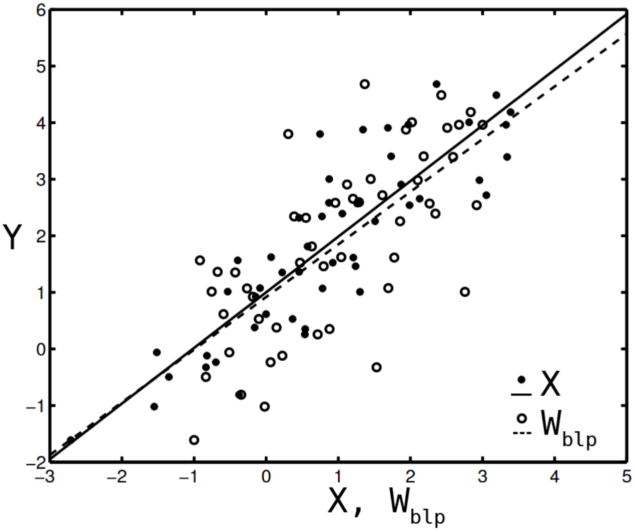
Regressing Y on W does not estimate the true slope β_X. Rather it estimates

$$Y = \beta_X^* W + \beta_0^* \quad (3)$$

In the *specific* case where X is normally distributed with mean μ_X and variance σ_X^2 it has been shown that

$$\beta_X^* = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_U^2}\beta_X = \lambda\beta_X \quad (4)$$

where σ_U^2 is the variance of the measurement error U.



[Carroll, 2006]

Define the "best linear predictor" of X given W to be

$$W_{blp} = (1 - \lambda)\mu_x + \lambda W \tag{5}$$

One can see that $X = W_{blp} + U^*$, where U^* is *uncorrelated with* X! Substituting into the true relationship between Y and X yields

$$Y = \beta_0 + \beta_X W_{blp} + \beta_X U^* \tag{6}$$

where the error $\beta_X U^*$ is *uncorrelated with* W_{blp}!
Thus, regressing Y on W_{blp} gives an unbiased estimate of the true relationship between Y and X.

There are two issues with naively adapting this procedure to our data:

- E_M is *lognormally* distributed in the solar wind and magnetosheath.
- The variation of ϕ_{PC} with E_M need not be linear.

Both of these issues can be resolved by doing analysis in logspace. Denoting the error-prone estimate of E_M from PRIME and PRIME-SH as E_M^*, we have the best linear predictor of E_M in logspace:

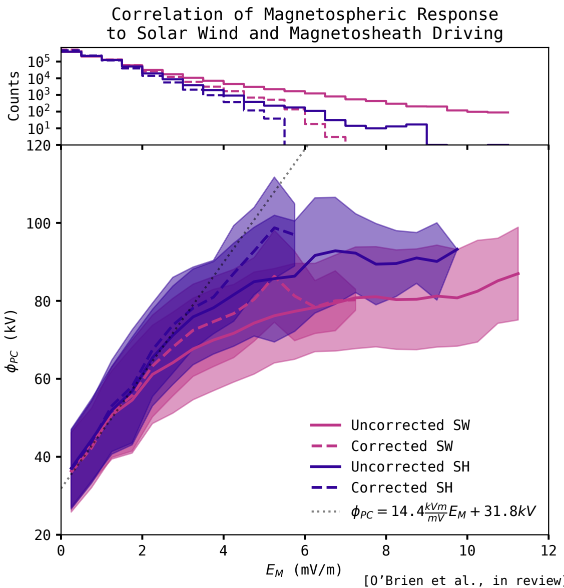
$$log(E_M)_{blp} = (1-\lambda)\mu_{log(E_M^*)} + \lambda log(E_M^*) \tag{7}$$

with $\lambda = \frac{\sigma^2_{E_M^*}}{\sigma^2_{E_M^*}+\sigma^2_U}$ calculable directly from PRIME and PRIME-SH's error estimates. Take the inverse log of each side to obtain the unbiased estimate of E_M:

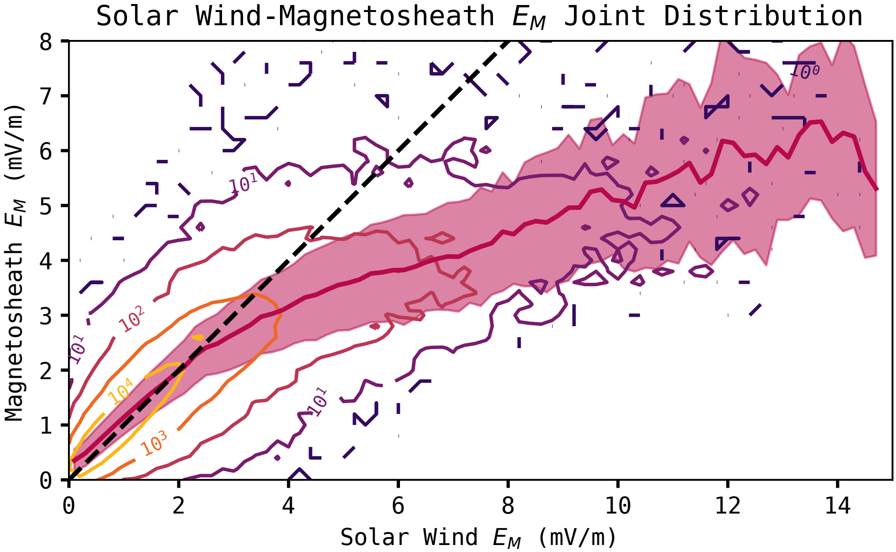
$$E_M = e^{(1-\lambda)\mu_{log(E_M^*)}} E_M^{*\lambda} \tag{8}$$

- ϕ_{PC} saturates less with the bias-corrected E_M in the solar wind and magnetosheath.
- ϕ_{PC} saturates with respect to the corrected solar wind $E_M > 4.5 mV/m$
- ϕ_{PC} does not appear to saturate with respect to the corrected magnetosheath E_M!

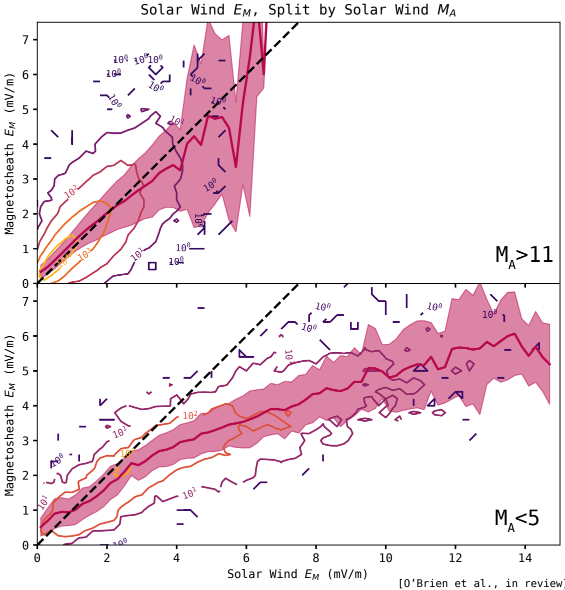
This implies there is some process in the magnetosheath reducing E_M relative to its value in the solar wind.



[O'Brien et al., in review]



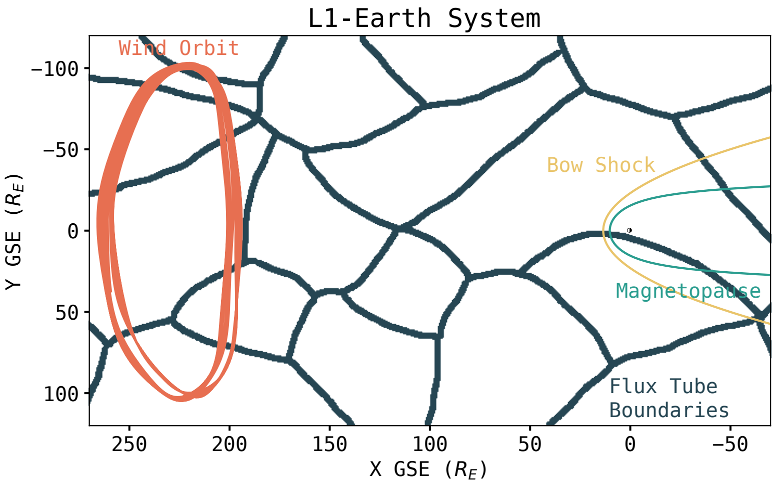
[O'Brien et al., in review]



[O'Brien et al., in review]

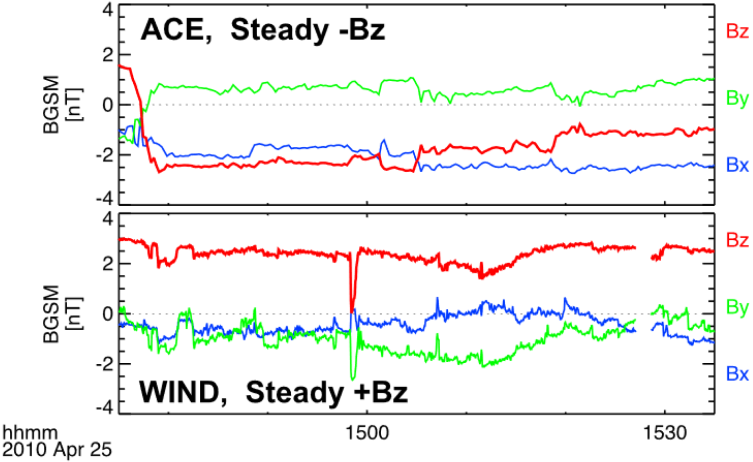
Conclusions:

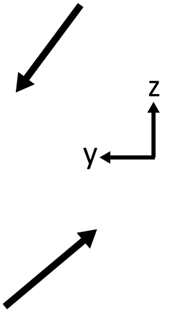
- Measurement uncertainty strongly affects statistical studies of solar wind-magnetosphere coupling.
- The ϕ_{PC} response to the magnetosheath electric field is linear after correcting for input uncertainty.
- Observed ϕ_{PC} saturation is associated with low solar wind M_A.



- The only continuous solar wind monitors are located at L1, 1,500,000km away from Earth
- There is no continuous monitor in the plasma actually contacting the magnetopause (the magnetosheath)

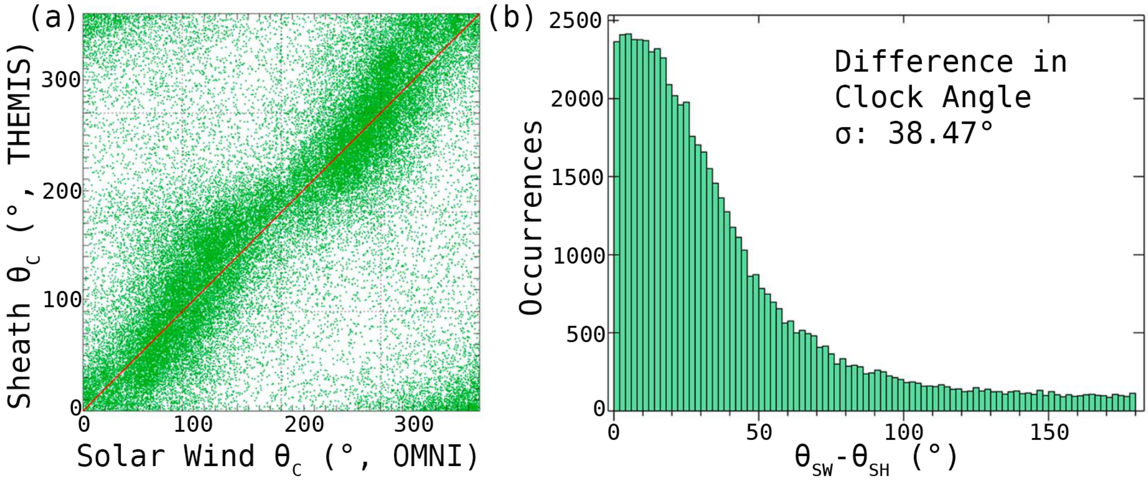
Solar Wind Uncertainty





Discrepancy in the interplanetary magnetic field between two L1 monitors (Walsh et al. 2019). Both monitors measure B_Z to be steady but of opposite sign.

Bow Shock Effects



Difference in IMF clock angle between solar wind (OMNI) and magnetosheath (THEMIS) measurements. (Walsh et al. 2019)

Continuous Rank Probability Score (CRPS) loss function:

$$CRPS = \int_{\infty}^{\infty} [C(y) - H(y - y_{obs})]^2 dy$$

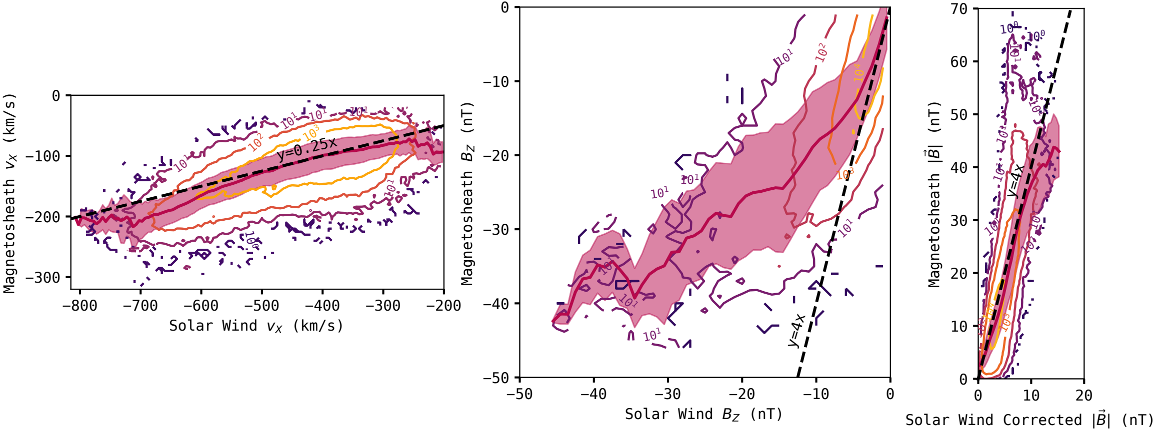
$C(y)$: Cumulative distribution $H(y - y_{obs})$: Heaviside function
For a Gaussian distribution with mean μ and variance σ^2 this becomes

$$CRPS(\mu, \sigma) = \sigma[\frac{y_{obs} - \mu}{\sigma} erf(\frac{(y_{obs} - \mu)}{\sqrt{2}\sigma}) + \sqrt{\frac{2}{\pi}} exp(-\frac{(y_{obs} - \mu)^2}{2\sigma^2}) - sqrt\frac{1}{\pi}]$$

This function is often better than the negative log likelihood for training models because it:

- Symmetrically punishes over- and underconfident predictions
- Collapses to the mean absolute error for $\sigma \to 0$
- Has the same unit as the variable of interest

Solar Wind-Magnetosheath v_x, B_Z, and $|\vec{B}|$ Joint Distributions



[O'Brien et al., in review]

The components of E_M along with the total magnetic field magnitude $|\vec{B}|$, with regression recalibration applied to correct for measurement uncertainty.