Machine learning based discovery of missing physical processes in radiation belt modeling

Enrico Camporeale (enrico.camporeale@noaa.gov)

CIRES / CU Boulder & NOAA Space Weather Prediction Center

Co-authors: G. Wilkie (PPPL), A. Drozdov, J. Bortnik (UCLA)

This project is supported by NASA under grant 80NSSC20K1580

Manuscript available on https://arxiv.org/abs/2107.14322

University of Colorado Boulder

CIRES

NOAA

UCLA

PPPL PRINCETON PLASMA PHYSICS LABORATORY

Inverse problem statement

$$\frac{\partial f(L,t)}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f(L,t)}{\partial L}\right) - \frac{f(L,t)}{\tau}$$

- What is the <u>optimal</u> choice of parameters (D_{LL} and τ) that makes the result of the diffusion equation most consistent with data?
- This is an INVERSE PROBLEM (we know the result, and want to infer the inputs), which is much harder than the “forward” model.

The Physics-Informed Neural Network (PINN) approach to parameter estimation

[HTML] **Physics-informed neural networks**: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations

M Raissi, P Perdikaris, GE Karniadakis - Journal of Computational **Physics**, 2019 - Elsevier

We introduce physics-informed neural networks–neural networks that are trained to solve supervised learning tasks while respecting any given laws of physics described by general nonlinear partial differential equations. In this work, we present our developments in the …

☆ 99 Cited by 1445 Related articles All 5 versions

PINN (Physics-Informed Neural Network) in a nuthsell

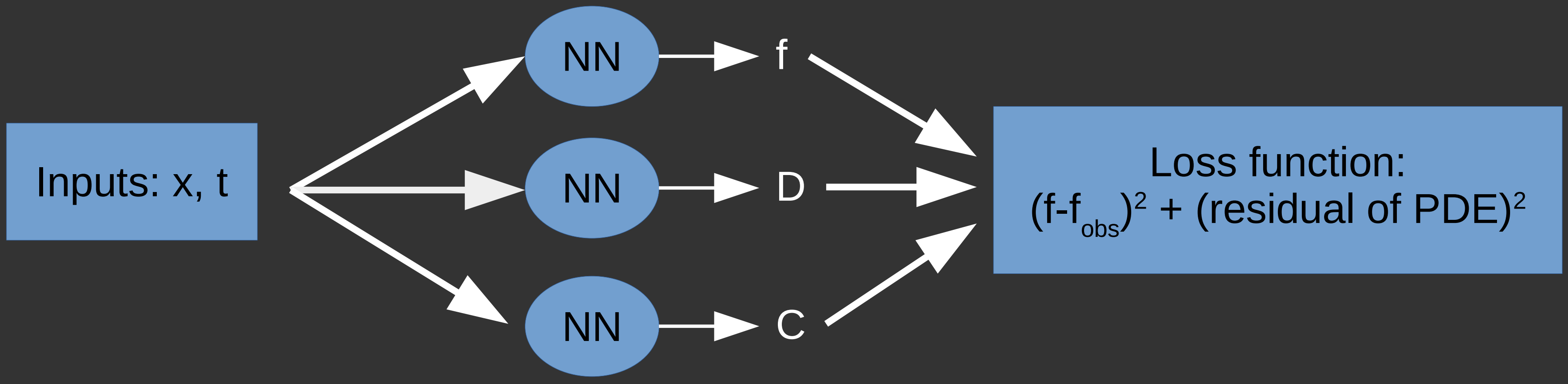
- PINN idea: to include the Partial Differential Equation (PDE) we want to solve in the cost function!

$$\mathcal{C}[f, D_{LL}, \tau] = \left[\frac{\partial f}{\partial t} - L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f}{\partial L}\right) + \frac{f}{\tau}\right]^2 + \left(f - f_{obs}\right)^2$$

- The trick under the hood: autodiff (automatic differentiation). All derivatives are computed exactly (using chain rule) !
- This is both:
  - a grid-less method to solve a PDE on a set of points (forward)
  - a way of estimating the coefficients of a PDE (inverse problem)

PINN for parameter estimation

- PDE: df(x,t) /dt + H[f(x,t);D,C] = 0
- (x,t) → (space,time)
- D,C, … → parameters



Clearing up some misconceptions

- How can you PREDICT the Phase Space Density simply as a function of space and time (L,t) ??

Clearing up some misconceptions

- How can you PREDICT the Phase Space Density simply as a function of space and time (L,t) ??

- I am not PREDICTING anything (not yet…)
- NN is a universal function approximator, which means that with enough complexity (i.e. number of weights), ANY function can be fitted “exactly” on a finite number of points.
- Among the infinite number of fitting functions, the PINN chooses one that also (approximately) solves the PDE on a set of points

Fokker-Planck equation

- Standard equation for radial diffusion:

$$\frac{\partial f(L,t)}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f(L,t)}{\partial L}\right) - \frac{f(L,t)}{\tau}$$

- If we allow D_{LL} and tau to be 'too general', this becomes ill-posed for an inverse problem (i.e., infinite combinations of D_{LL} and tau give the same result)
- Here we use the drift-diffusion form of FP equation:

$$\frac{\partial f(L,t)}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f(L,t)}{\partial L}\right) - \frac{\partial C f(L,t)}{\partial L},$$

Diffusion coefficient (D_{LL})

Drift coefficient (C)

Baseline models

$$D_{LL}^{BA} = L^{10} \cdot 10^{(0.506Kp - 9.325)}$$

$$D_{LL}^{Ozeke} = 2.6 \cdot L^{6} \cdot 10^{(0.217L + 0.461Kp - 8)} + 6.62 \cdot L^{8} \cdot 10^{(-0.0327L^2 + 0.625L - 0.0108Kp^2 + 0.499Kp - 13)}$$

Brautigam & Albert *JGR* (2000)
Ozeke et al. *JGR* (2014)

$$\tau = 10 \text{ for } L \leq L_{pp}$$
$$= 6/Kp \text{ for } L > L_{pp}$$

Shprits et al. *Annales Geophys.* (2005)
Drozdov et al. *Space Weather* (2016)

Errors defined in Morley et al. *Space Weather* (2018)

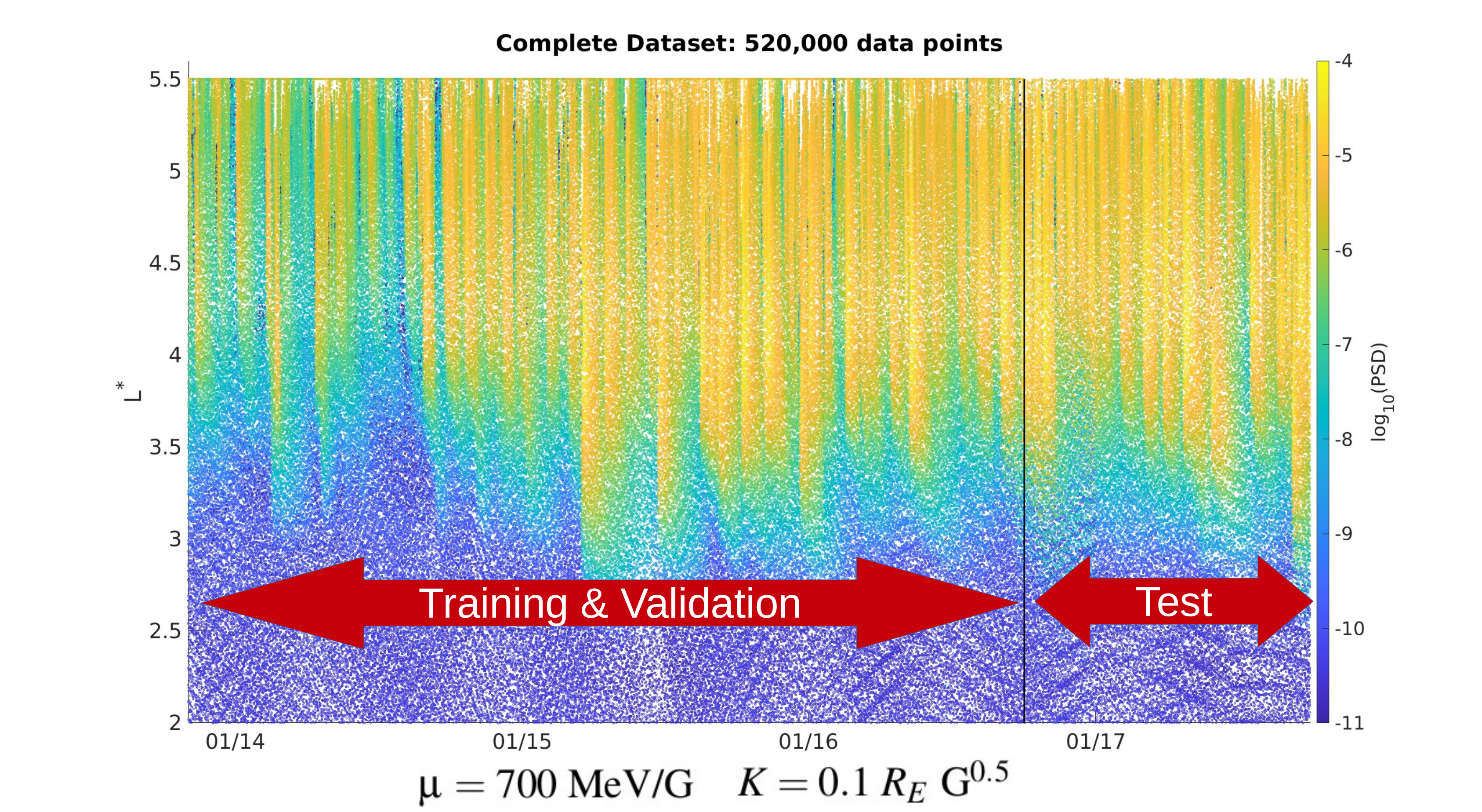
Plasmapause location is estimated using the approximation of *Carpenter and Anderson (1992)*

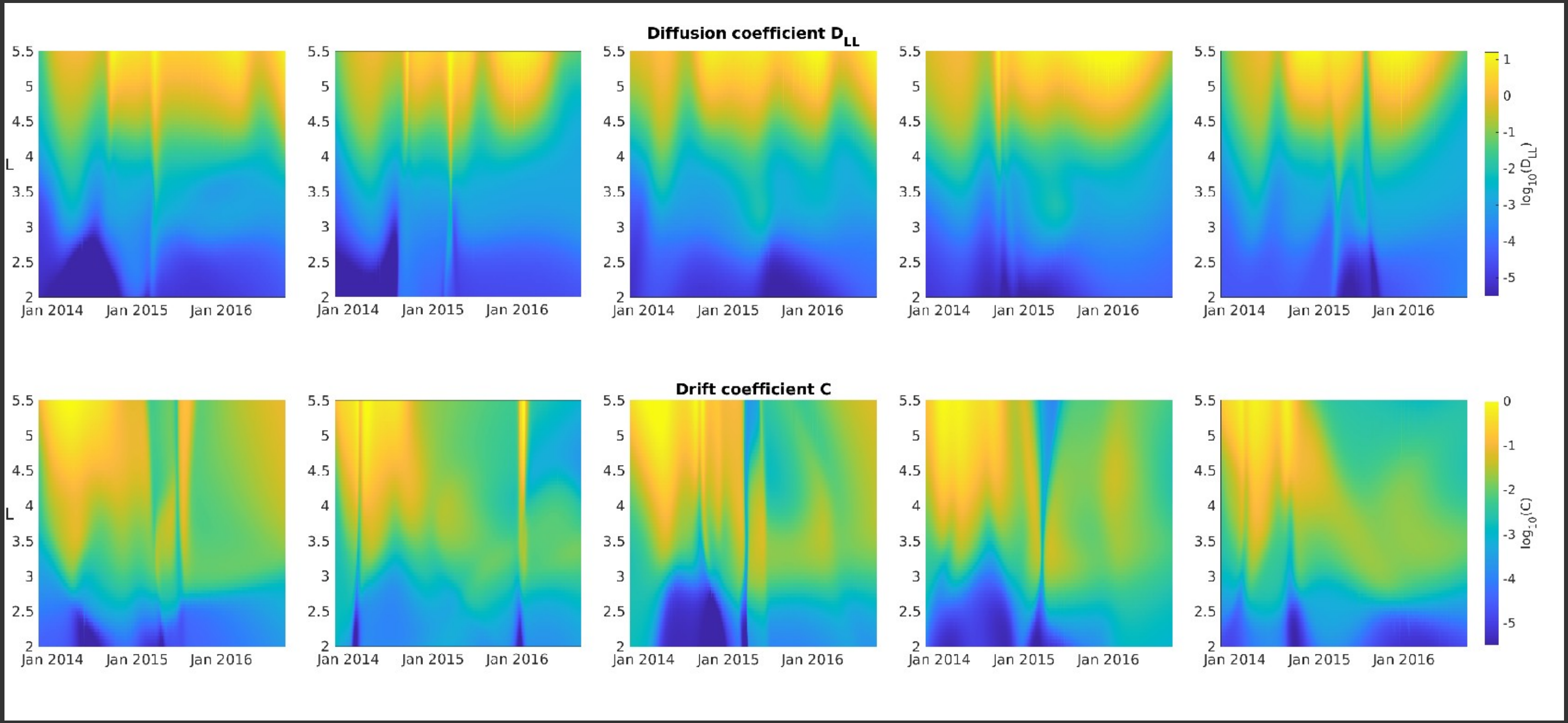
Percentage symmetric accuracy $$\zeta_k = 100 \cdot \exp(P_k(|\log(f/\hat{f})|)),$$

Symmetric signed percentage bias $$\text{SSPB} = 100 \cdot \text{sgn}(P_{50}(\log(f/\hat{f})))(\exp(|P_{50}(\log(f/\hat{f}))|) - 1)$$

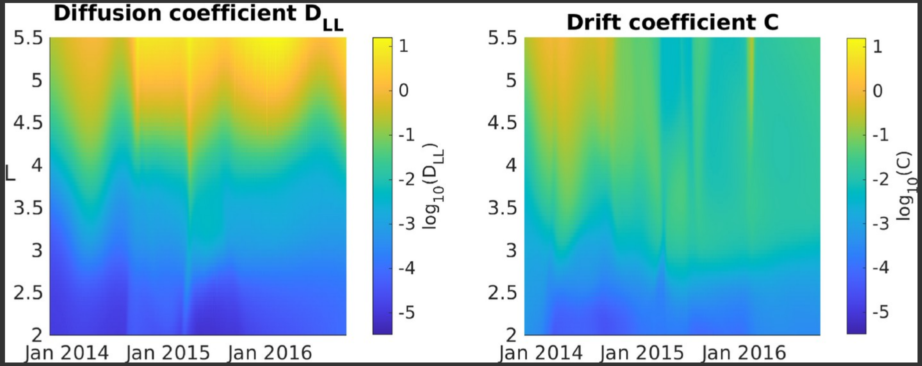
Unconditionally stable forward model (2nd order 'modified' Crank-Nicolson finite difference scheme) as in Welling, D. T., Koller, J., & Camporeale, E. (2012). Verification of SpacePy's radial diffusion radiation belt model. *Geoscientific Model Development*, 5(2), 277-287.

Van Allen Probes data



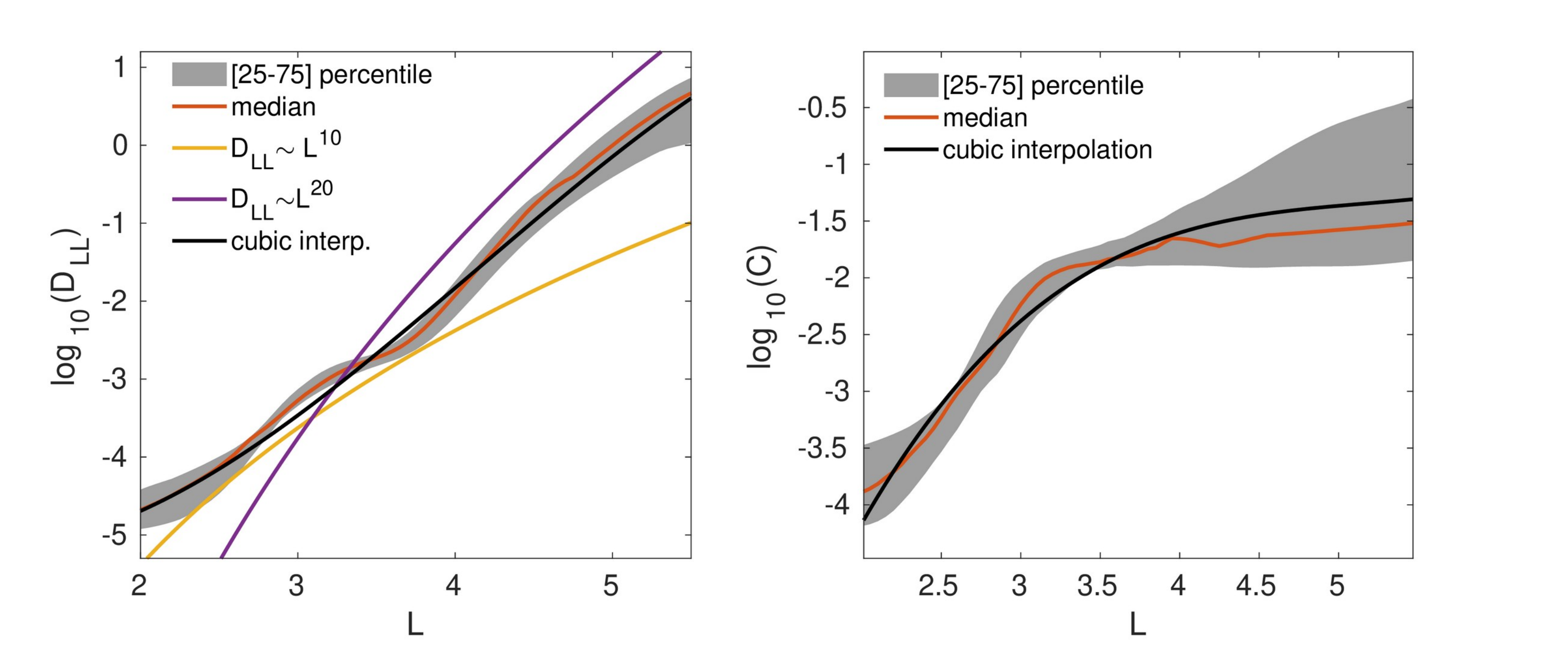


"Best" 5 solutions in an ensemble of 20



Average of the best 5 solutions

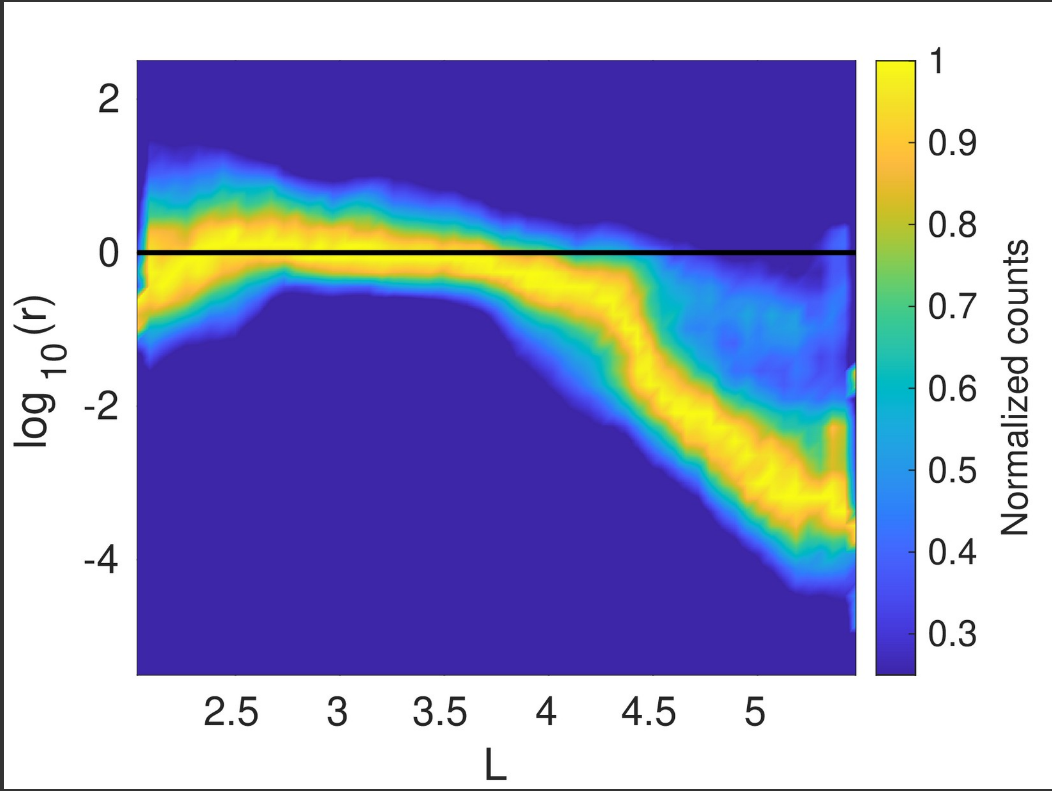
Statistical analysis of coefficients



$$\log_{10} D_{LL} = -0.0593L^3 + 0.7368L^2 - 1.33L - 4.505$$

$$\log_{10} C = 0.0777L^3 - 1.2022L^2 + 6.3177L - 12.6115$$

Relative importance between drift and diffusion

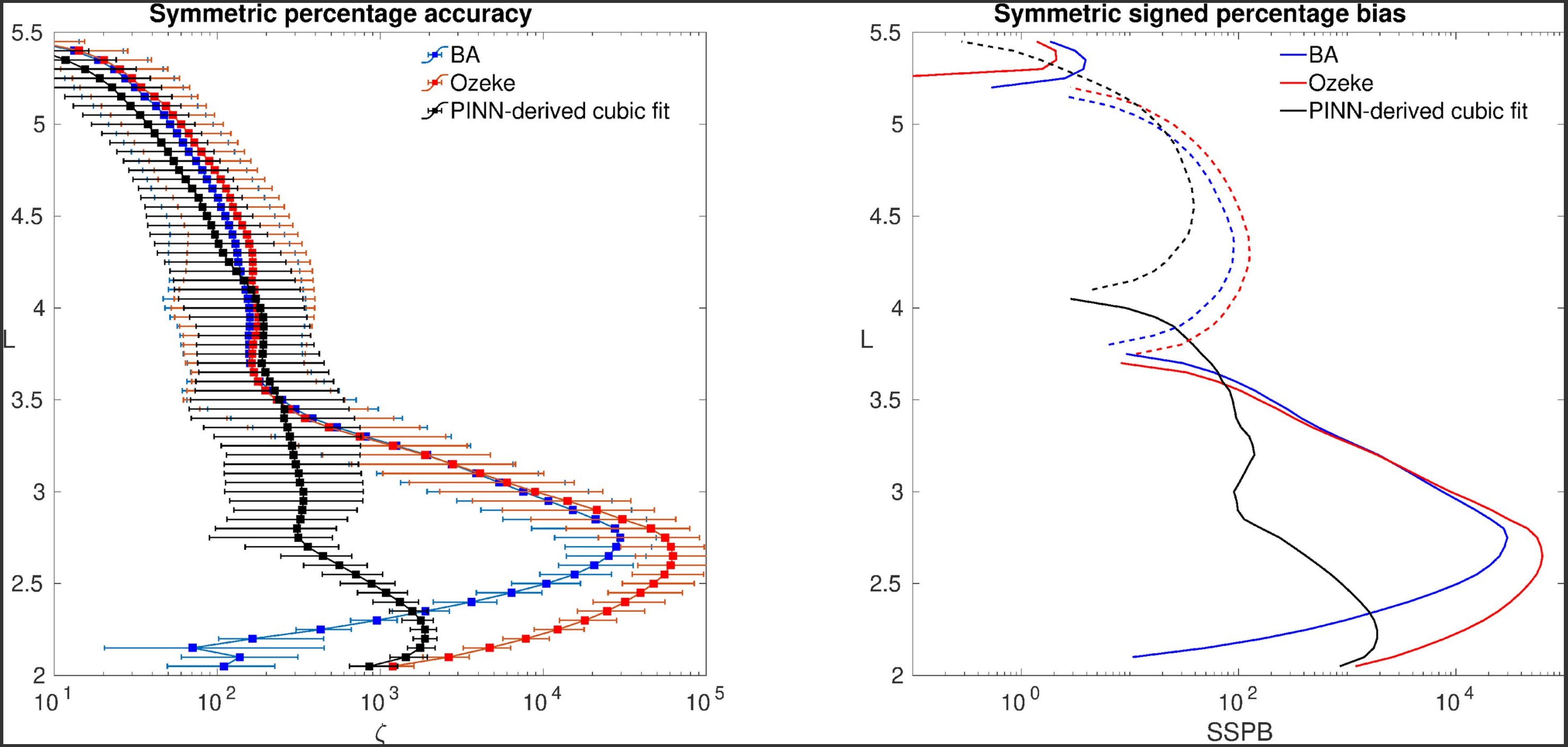


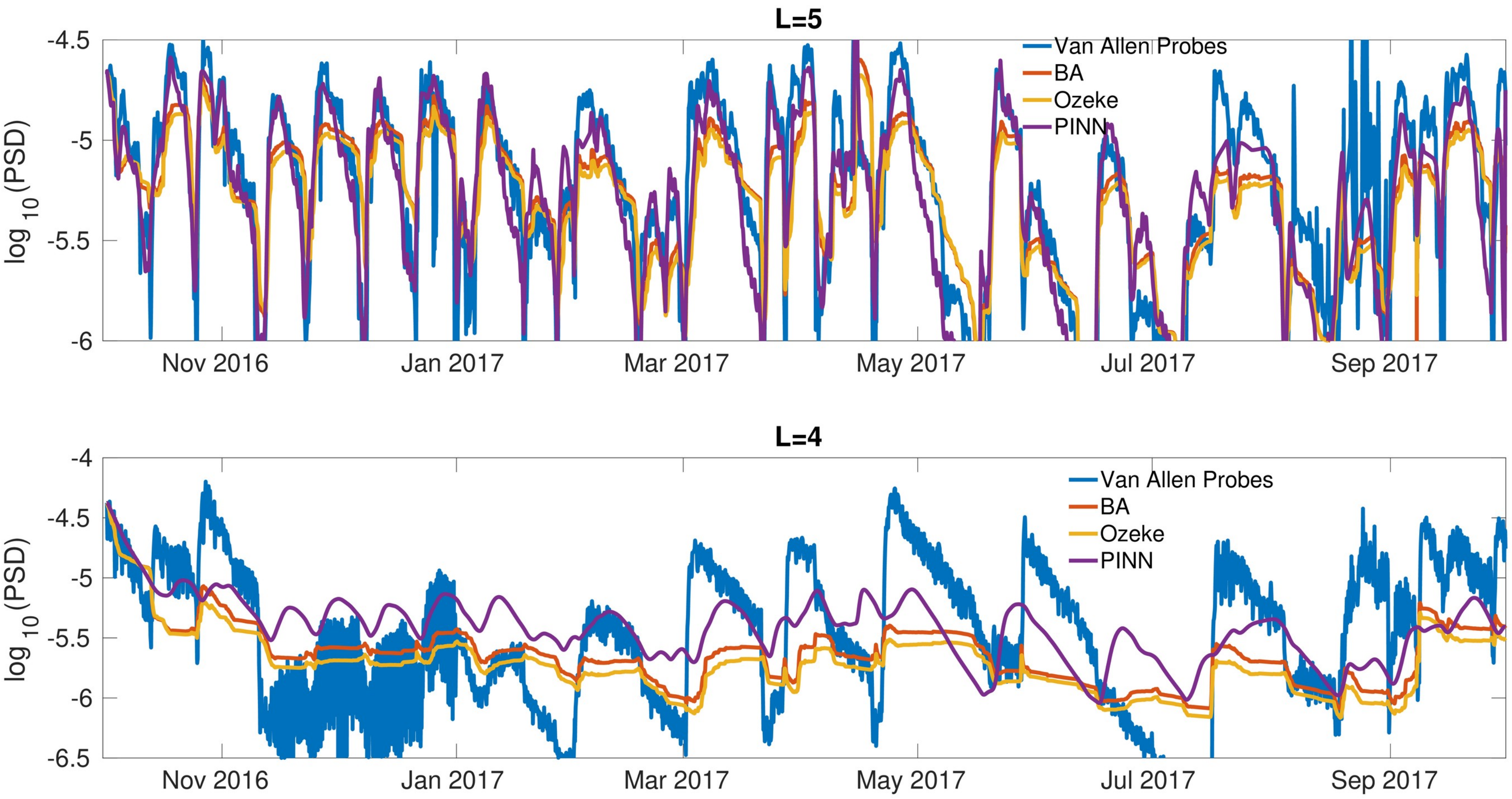
$$r = \left| \frac{1}{L^2} \left(\frac{\partial C f}{\partial L} \right) / \left[\frac{\partial}{\partial L} \left(\frac{D_{LL}}{L^2} \frac{\partial f}{\partial L} \right) \right] \right|$$

r = Drift term over diffusion term

Drift and diffusion
are comparable for L<4

Results on test set when using PINN-learned coefficients (cubic interpolation)




L=5
-4.5
-5
-5.5
-6
log 10(PSD)
Nov 2016
Jan 2017
Mar 2017
May 2017
Jul 2017
Sep 2017
Van Allen Probes
BA
Ozeke
PINN
L=4
-4
-4.5
-5
-5.5
-6
-6.5
log 10(PSD)
Nov 2016
Jan 2017
Mar 2017
May 2017
Jul 2017
Sep 2017
Van Allen Probes
BA
Ozeke
PINN

Conclusions

- Physics-informed Neural Network (PINN) is a **game changer** in applied mathematics
- Possibly, the best way we can use it in space physics problems is to solve **inverse problems**
- Use case: derivation of Diffusion and Drift coefficients for radial transport for electrons in radiation belt
- With PINN you can:
  - Re-parametrize the diffusion and drift coefficients with simple, interpretable relationships (preprint: https://arxiv.org/abs/2107.14322)
  - Assess when quasi-linear assumptions are not valid
  - Automatically identify interesting events
  - Re-define electron lifetimes
  - Improve accuracy in nowcasting/forecasting (without having to use future Kp values)

ML-Helio 2022

Abstract submission is open on
https://ml-helio.github.io/

Contact: Enrico.camporeale@noaa.gov