Using Physics-Informed Neural Networks for solving inverse problems: a space physics case study

Enrico Camporeale

CIRES / CU Boulder & NOAA Space Weather Prediction Center

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

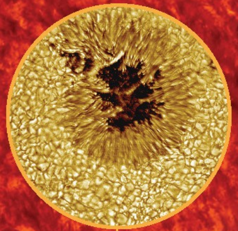
This project is supported by NASA under grant 80NSSC20K1580

Space Weather

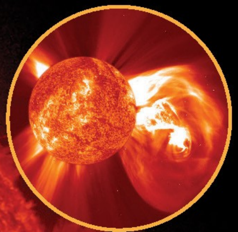
Space weather refers to the variable conditions on the Sun and in the space environment that can influence the performance and reliability of space-based and ground-based technological systems, as well as endanger life or health. Just like weather on Earth, space weather has its seasons, with solar activity rising and falling over an approximate 11 year cycle.

Sunspots

Sunspots are comparatively cool areas at up to 7,700° F and show the location of strong magnetic fields protruding through what we would see as the Sun's surface. Large, complex sunspot groups are generally the source of significant space weather.



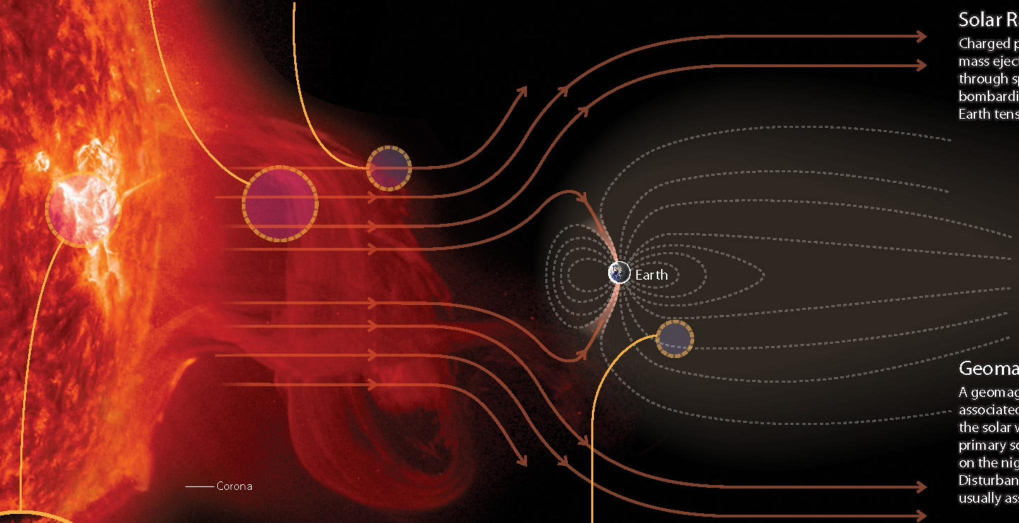
Coronal Mass Ejections (CMEs)



Large portions of the corona, or outer atmosphere of the Sun, can be explosively blown into space, sending billions of tons of plasma, or superheated gas, Earth's direction. These CMEs have their own magnetic field and can slam into and interact with Earth's magnetic field, resulting in geomagnetic storms. The fastest of these CMEs can reach Earth in under a day, with the slowest taking 4 or 5 days to reach Earth.

Solar Wind

The solar wind is a constant outflow of electrons and protons from the Sun, always present and buffeting Earth's magnetic field. The background solar wind flows at approximately one million miles per hour!

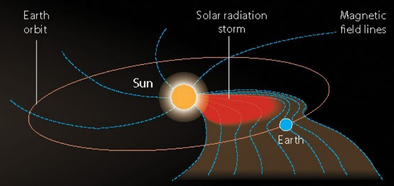


Sun's Magnetic Field

Strong and ever-changing magnetic fields drive the life of the Sun and underlie sunspots. These strong magnetic fields are the energy source for space weather and their twisting, shearing, and reconnection lead to solar flares.

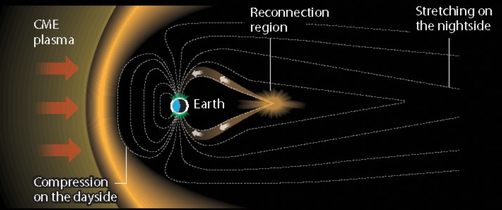
Solar Radiation Storms

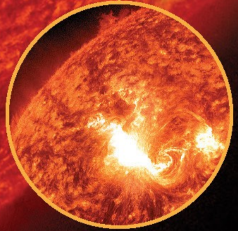
Charged particles, including electrons and protons, can be accelerated by coronal mass ejections and solar flares. These particles bounce and gyrate their way through space, roughly following the magnetic field lines and ultimately bombarding Earth from every direction. The fastest of these particles can affect Earth tens of minutes after a solar flare.



Geomagnetic Storms

A geomagnetic storm is a temporary disturbance of Earth's magnetic field typically associated with enhancements in the solar wind. These storms are created when the solar wind and its magnetic field interacts with Earth's magnetic field. The primary source of geomagnetic storms is CMEs which stretch the magnetosphere on the nightside causing it to release energy through magnetic reconnection. Disturbances in the ionosphere (a region of Earth's upper atomosphere) are usually associated with geomagnetic storms.





Solar Flares

Reconnection of the magnetic fields on the surface of the Sun drive the biggest explosions in our solar system. These solar flares release immense amounts of energy and result in electromagnetic emissions spanning the spectrum from gamma rays to radio waves. Traveling at the speed of light, these emissions make the 93 million mile trip to Earth in just 8 minutes.

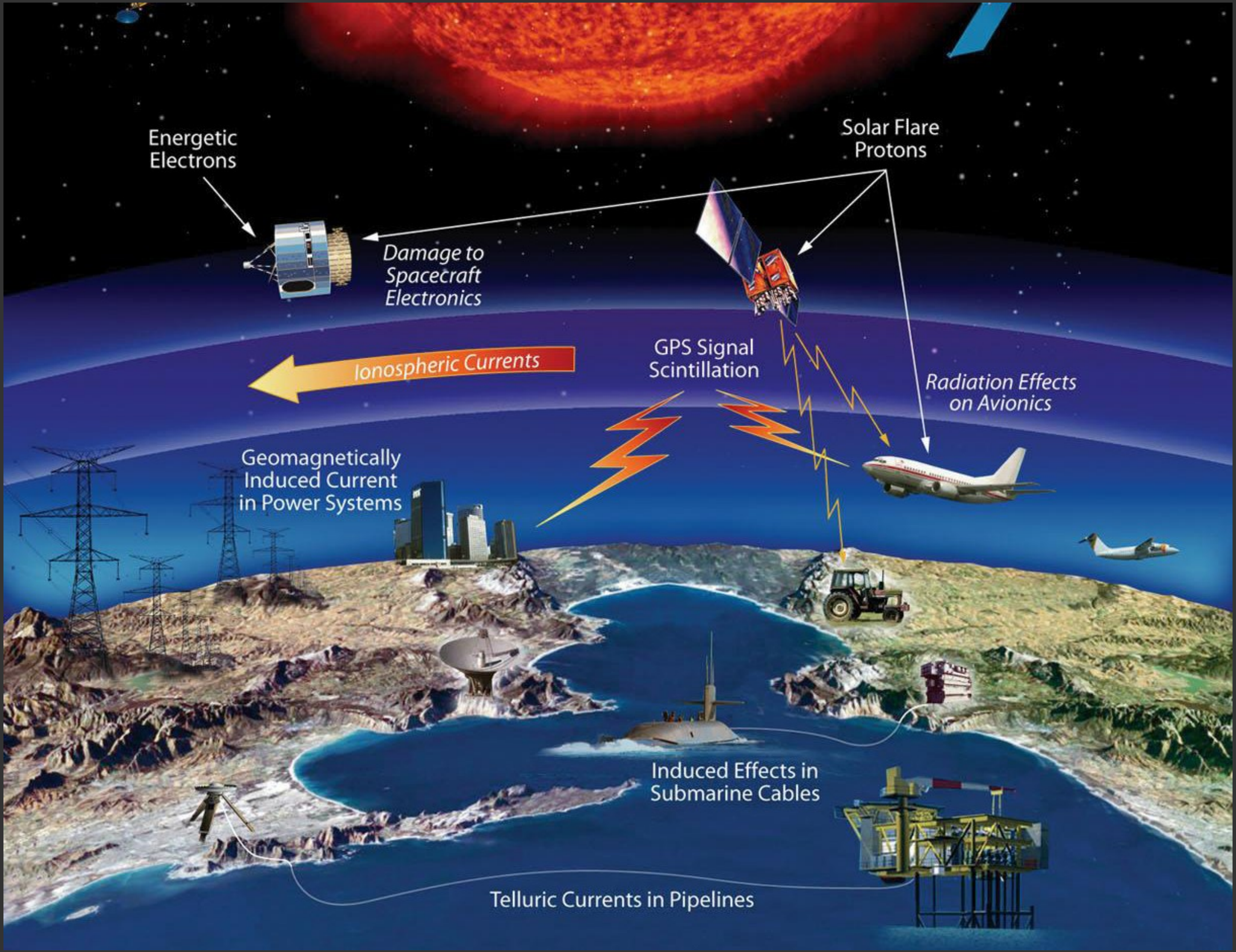
Earth's Magnetic Field

Earth's magnetic field, largely like that of a bar magnet, gives the Earth some protection from the effects of the Sun. Earth's magnetic field is constantly compressed on the day side and stretched on the night side by the ever-present solar wind. During geomagnetic storms, the disturbances to Earth's magnetic field can become extreme. In addition to some buffering by the atmosphere, this field also offers some shielding from the charged particles of a radiation storm.

NOAA

NOAA Space Weather Prediction Center – www.spaceweather.gov

Source images: NASA, NOAA.

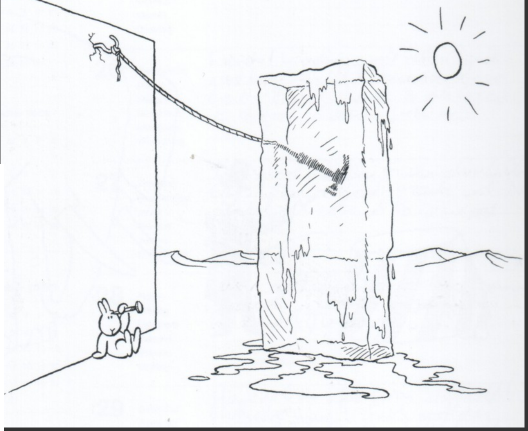


Credits: NASA

A disaster waiting to happen…

- If a space weather event causes a power outage, we estimate costs to U.S. electricity consumers that may be ~$400 million to ~$10 billion for a moderate event and ~$1 billion to ~$20 billion for a more extreme event.
- 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**

The book of bunny suicides



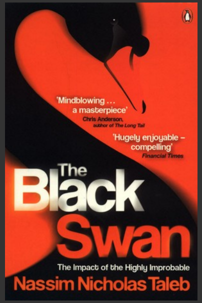

FINAL REPORT

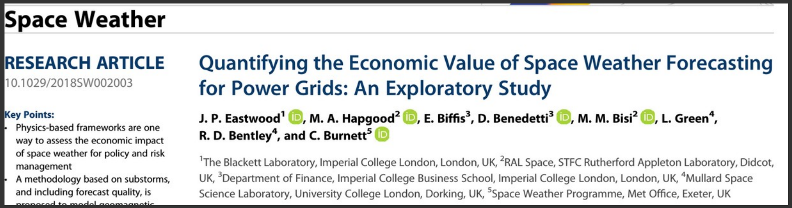
Social and Economic Impacts of Space Weather in the United States

September 2017

Abt Associates
Bethesda, Maryland

Written under contract for the NOAA National Weather Service




Space Weather

RESEARCH ARTICLE
10.1029/2018SW002003

Quantifying the Economic Value of Space Weather Forecasting for Power Grids: An Exploratory Study

J. P. Eastwood[1], M. A. Hapgood[2], E. Biffis[3], D. Benedetti[3], M. M. Bisi[2], L. Green[4], R. D. Bentley[4], and C. Burnett[5]

[1]The Blackett Laboratory, Imperial College London, London, UK, [2]RAL Space, STFC Rutherford Appleton Laboratory, Didcot, UK, [3]Department of Finance, Imperial College Business School, Imperial College London, London, UK, [4]Mullard Space Science Laboratory, University College London, Dorking, UK, [5]Space Weather Programme, Met Office, Exeter, UK

Key Points:
- Physics-based frameworks are one way to assess the economic impact of space weather for policy and risk management
- A methodology based on substorms, and including forecast quality, is

A disaster waiting to happen…
and it does happen!

CNN BUSINESS Markets **Tech** Media Success Perspectives Videos LIVE TV CNN+

SpaceX will lose up to 40 satellites it just launched due to a solar storm

By Jackie Wattles, CNN Business

Updated 7:44 PM ET, Wed February 9, 2022

SPACE.com

News Tech Spaceflight Science & Astronomy Search For Life

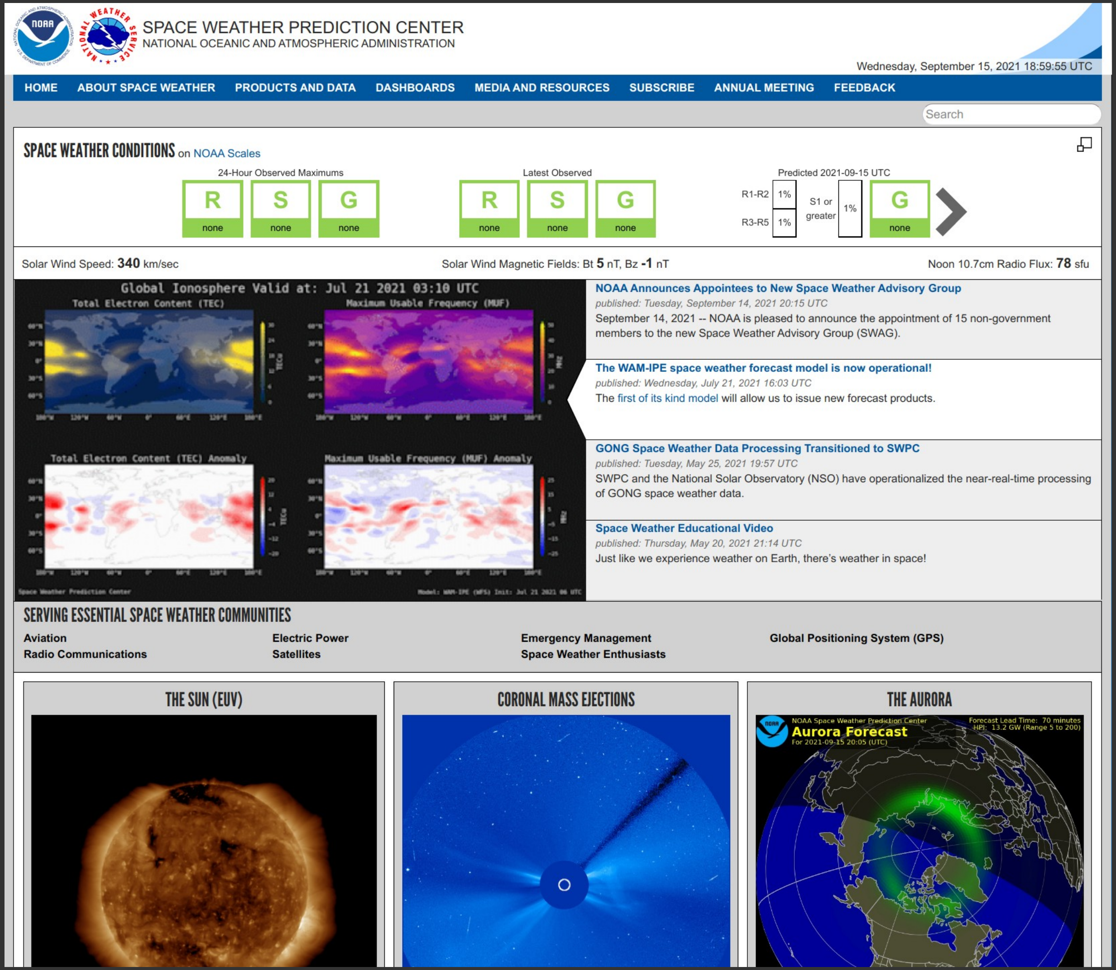
Home > News > Spaceflight

Solar geomagnetic storms could threaten more satellites after Elon Musk's Starlink

By Chelsea Gohd published 28 days ago

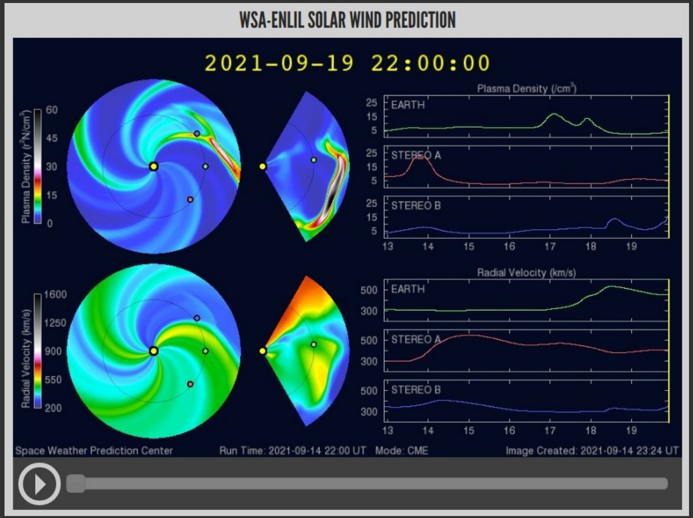
"That is a drag," NOAA's Bill Murtagh said.

NOAA Space Weather Prediction Center

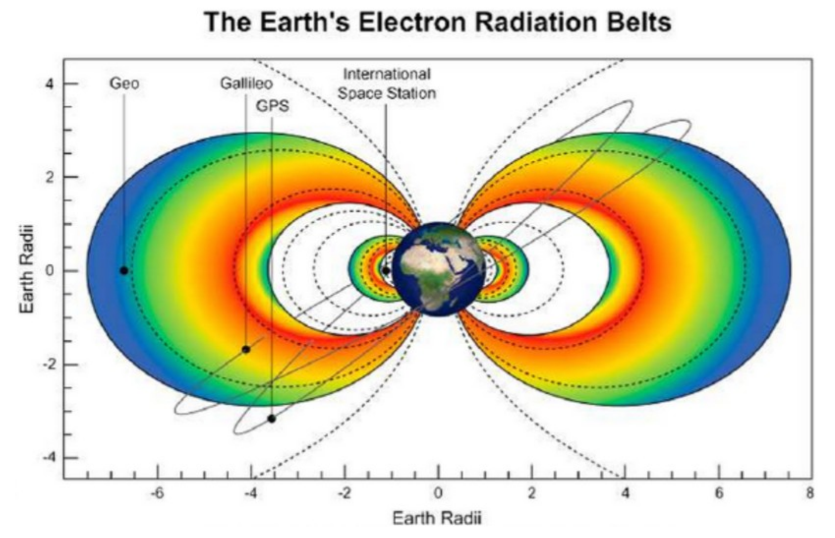


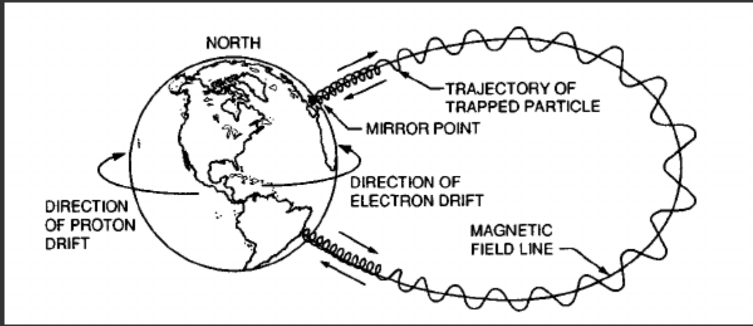
SWPC mission:

Safeguarding society with actionable space weather information



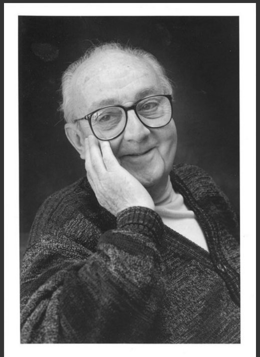
Radiation belts in one slide





Adiabatic invariants are approximate constants associated to a periodic orbit: their values are constant provided the forces vary at much slower rate than the corresponding period.

Orbit		Adiabatic invariant
Gyromotion	→	Magnetic moment (μ, 1st adiabatic inv.)
Bounce orbit	→	Integral invariant (J, 2nd adiabatic inv.)
Drift orbit	→	Flux invariant (Φ, 3rd adiabatic inv. Related to L)



“All models are wrong, some are useful” (G. Box)

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

$f \rightarrow$ Phase Space Density: unit of flux/momentum – $(c/MeV/cm)^3$
$L \rightarrow$ (“L star”) Equatorial radial distance (3rd adiabatic invariant). Unit of R_E
$t \rightarrow$ time (unit of day)
$D_{LL} \rightarrow$ Diffusion coefficient (unit of L^2/t)
$\tau \rightarrow$ electron lifetime (unit of time)

The radial diffusion equation is intended for fixed values of 1st and 2nd adiabatic invariants.

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 optimal 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.

Bayesian parameter estimation

- The ‘standard’ statistical approach:

  Because the underlying model is “wrong” there is not a unique solution, rather we are looking for a **distribution of parameters**

- Classical topic in geoscience, but maybe not yet a mainstream approach in space physics!
- Difficult to apply in higher dimensions
- It typically still requires a parameterization of coefficients.

JGR Space Physics

RESEARCH ARTICLE

10.1029/2019JA027618

Key Points:

- We present the first application of Bayesian parameter estimation to the problem of quasi-linear radial diffusion in the radiation belt
- The Bayesian approach allows the problem to be cast in probabilistic terms and for ensemble simulations to be run
- An improved accuracy is demonstrated when compared against standard deterministic models

Bayesian Inference of Quasi-Linear Radial Diffusion Parameters using Van Allen Probes

Rakesh Sarma[1], Mandar Chandorkar[1], Irina Zhelavskaya[2,3], Yuri Shprits[2,3], Alexander Drozdov[4], and Enrico Camporeale[1,5,6]

[1]Centrum Wiskunde & Informatica, Amsterdam, The Netherlands, [2]GFZ German Research Centre For Geosciences, Potsdam, Germany, [3]Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany, Department of Earth, Planetary, and Space Sciences, [4]University of California, Los Angeles, CA, USA, [5]CIRES, University of Colorado, Boulder, CO, USA, [6]NOAA, Space Weather Prediction Center, Boulder, CO, USA

Abstract The Van Allen radiation belts in the magnetosphere have been extensively studied using

A different (non-Bayesian) 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 …

☆ Save 99 Cite Cited by 2559 Related articles All 6 versions Web of Science: 1014

PINN (Physics-Informed Neural Network) in a nuthsell

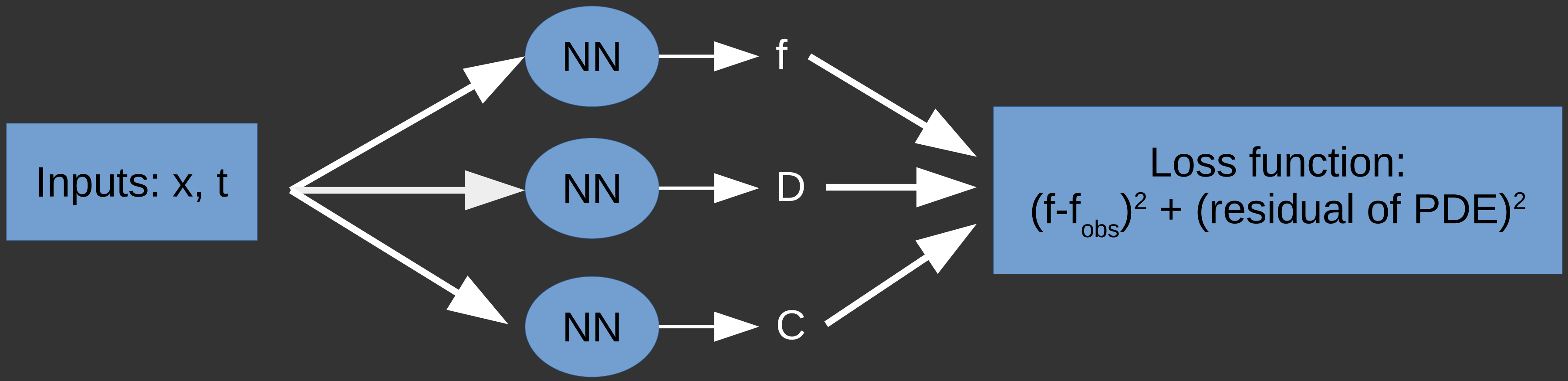
- PINN idea: to include the Partial Differential Equation (PDE) we want to solve in the cost function!

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

- The trick under the hood: autodiff (automatic differentiation). All derivatives are computed exactly (using chain rule) !
  - This can be done because the output of a NN is an analytical and differentiable function
- 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: $df(x,t)/dt + H[f(x,t);D,C] = 0$
- $(x,t) \rightarrow$ (space,time)
- D,C, … $\rightarrow$ parameters (functions of time and space)



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:

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

Drift coefficient (C)

Diffusion coefficient (D_{LL})

Baseline models

$$D_{LL}^{BA} = L^{10} \cdot 10^{(0.506Kp-9.325)}$$
$$D_{LL}^{Ozeke} = 2.6 \cdot L^6 \cdot 10^{(0.217L+0.461Kp-8)} + 6.62 \cdot L^8 \cdot 10^{(-0.0327L^2+0.625L-0.0108Kp^2+0.499Kp-13)}$$
$$D_{LL}^{Ali} = \exp(-16.951+0.181Kp \cdot L+1.982L)+\exp(-16.253+0.224Kp \cdot L+L)$$

Brautigam & Albert *JGR* (2000)
Ozeke et al. *JGR* (2014)
Ali et al. *JGR* (2016)

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

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

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

Definition of errors

Percentage Symmetric Accuracy:

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

where $\hat{f}$ and f are the ground-truth values taken by observations and the corresponding values produced by a model, respectively. P_k represents the $k-$th percentile (i.e. P_{50} is the median) calculated over all values at fixed L. This represents a general-

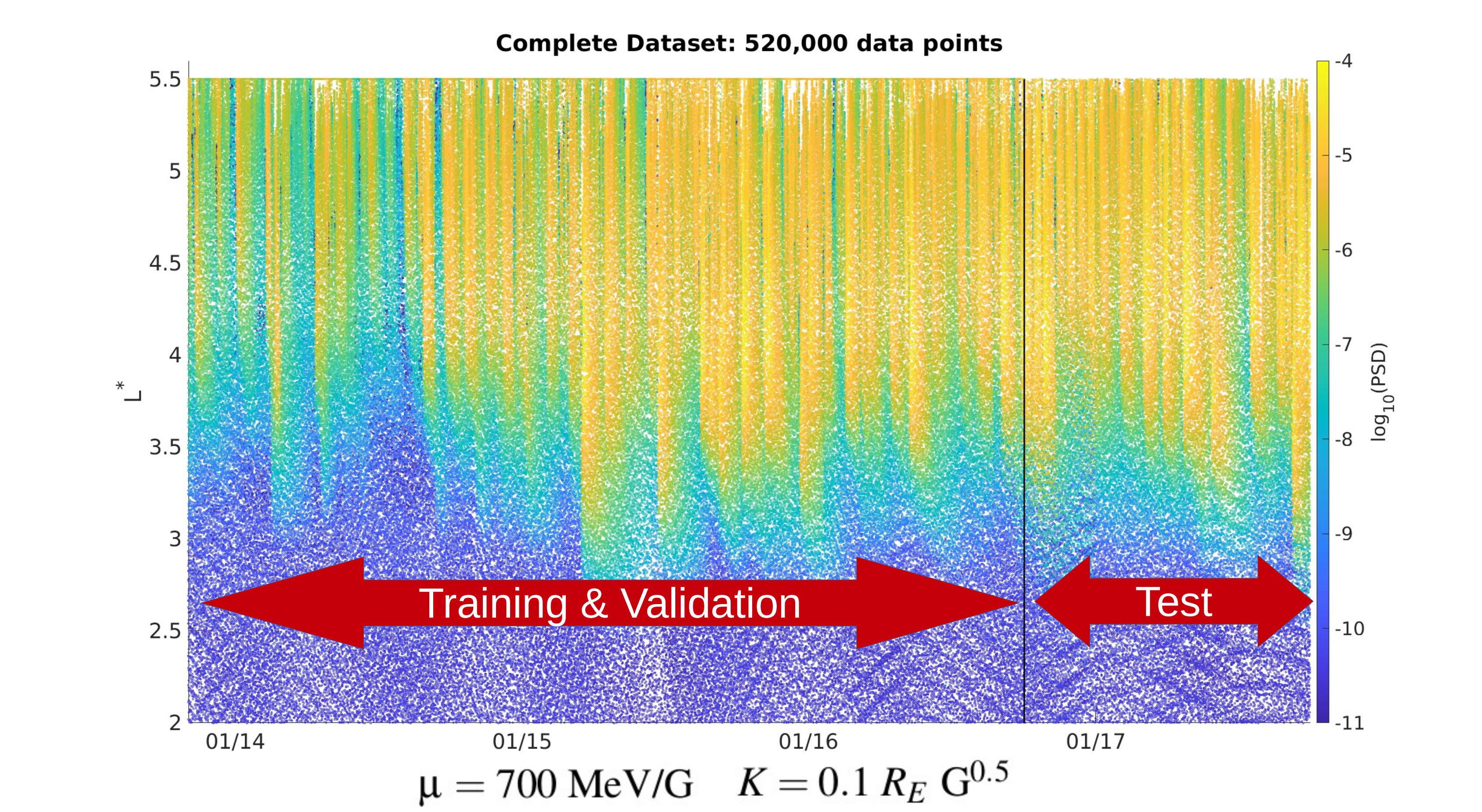
Symmetric Signed Percentage Bias:

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

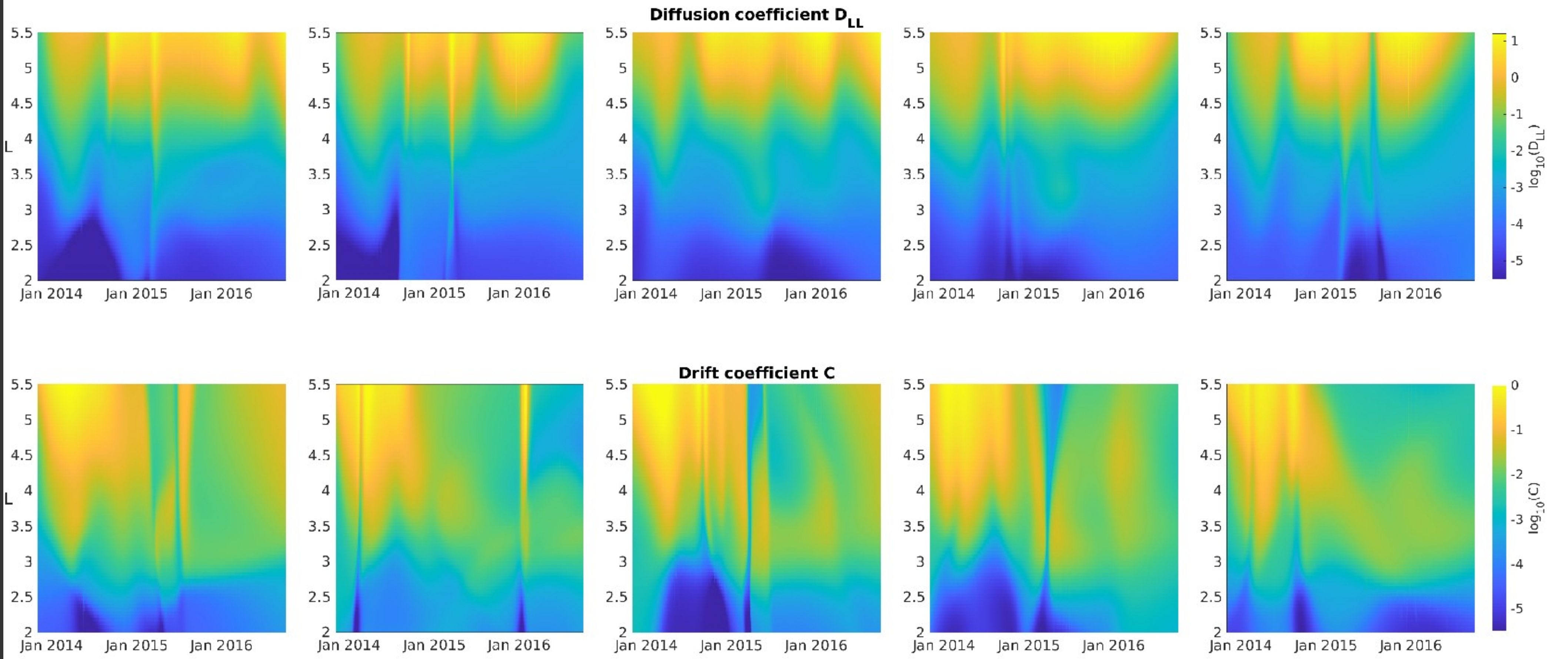
Relative Error of the log of Phase Space Density:

$$\varepsilon(L) = P_{50}\left(\frac{\log_{10} f - \log_{10} \hat{f}}{\log_{10} \hat{f}}\right)$$

Van Allen Probes data

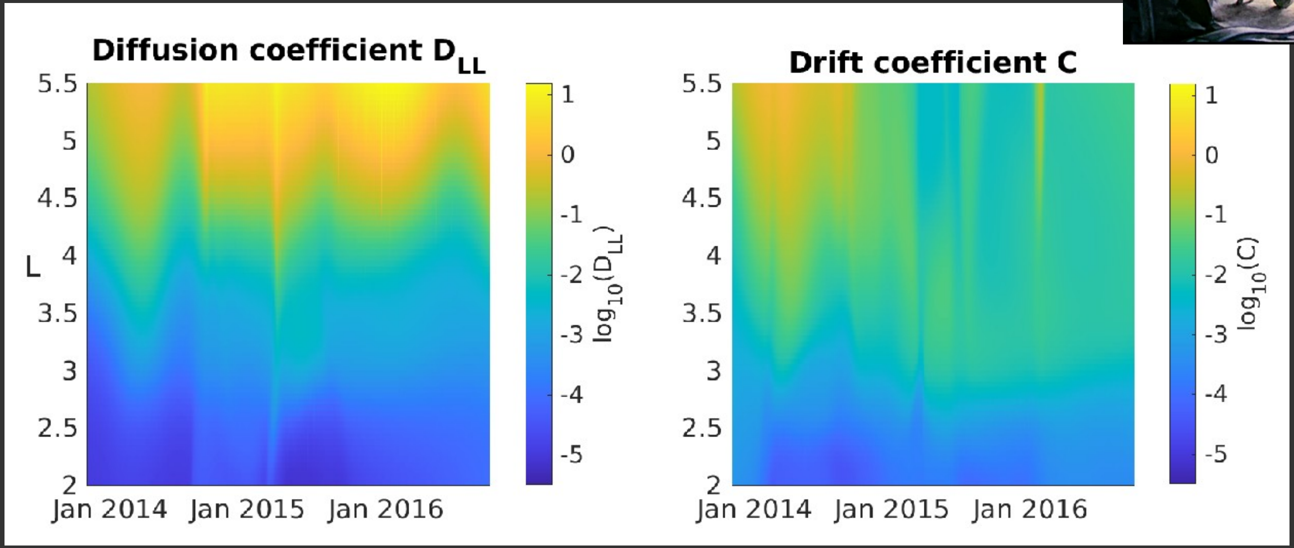


“Best” 5 solutions in an ensemble of 20



Data mining

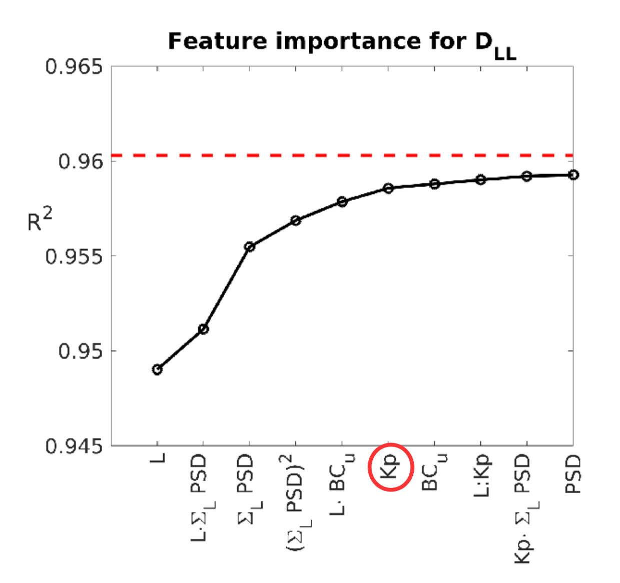
$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f}{\partial L}\right) - \frac{\partial}{\partial L}(Cf)$$

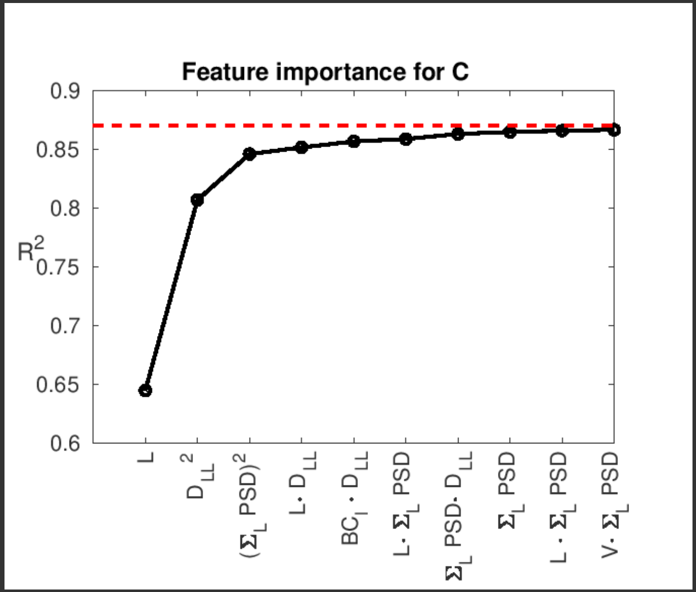


All the information about the physics of interest is in these two coefficients!

Feature selection

- Feature selection with backward elimination method (generalized linear model)





Can we forecast unseen data?

- We train an ensemble of decision trees with gradient boosting to predict the coefficients D_{LL} and C
  - Inputs: L, log10(PSD), running average of PSD over L, Boundary conditions
  - NO Kp!! (Different from all other parameterizations in the literature)
- Using PSD as input makes the problem non-linear: the Diffusion and drift coefficients are functions of the solution.
- Numerically $\rightarrow$ Predictor-Corrector method

Parameter estimation pipeline

Assumption:
The physics obeys FP equation

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

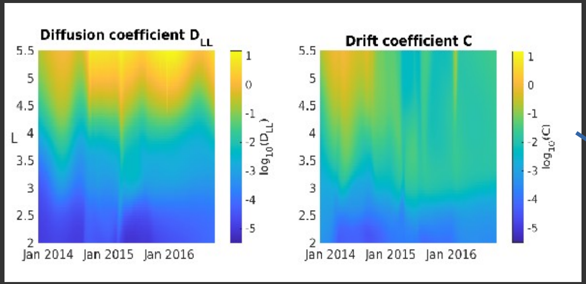
Parameter estimation pipeline

Assumption:
The physics obeys FP equation

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

PINN

Optimal
coefficients



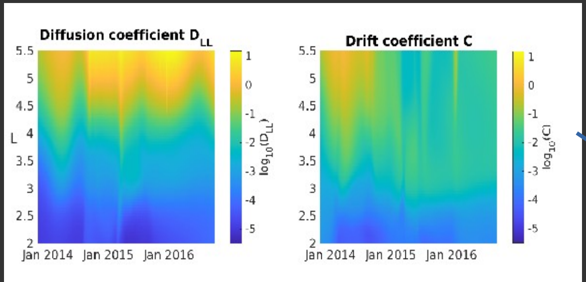
Parameter estimation pipeline

Assumption:
The physics obeys FP equation

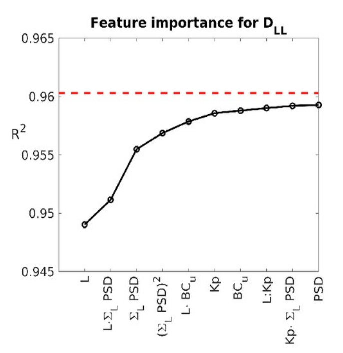
$$\frac{\partial f(L,t)}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f(L,t)}{\partial L}\right) - \frac{\partial C f(L,t)}{\partial L},$$

PINN

Optimal coefficients



Feature selection



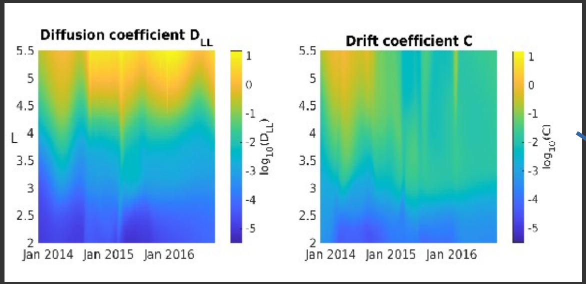
Parameter estimation pipeline

Assumption:
The physics obeys FP equation

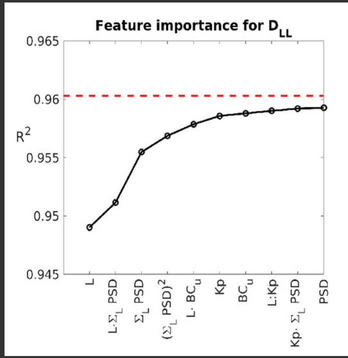
$$\frac{\partial f(L,t)}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f(L,t)}{\partial L}\right) - \frac{\partial C f(L,t)}{\partial L},$$

PINN

Optimal coefficients



Feature selection



Train a ML model that predicts D_{LL} and C at a given time

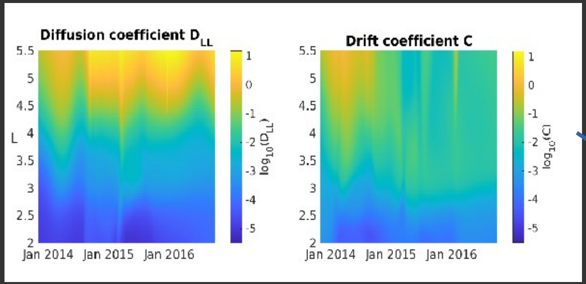
Parameter estimation pipeline

Assumption:
The physics obeys FP equation

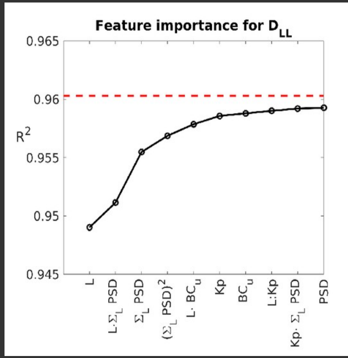
$$\frac{\partial f(L,t)}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f(L,t)}{\partial L}\right) - \frac{\partial C f(L,t)}{\partial L},$$

PINN

Optimal coefficients



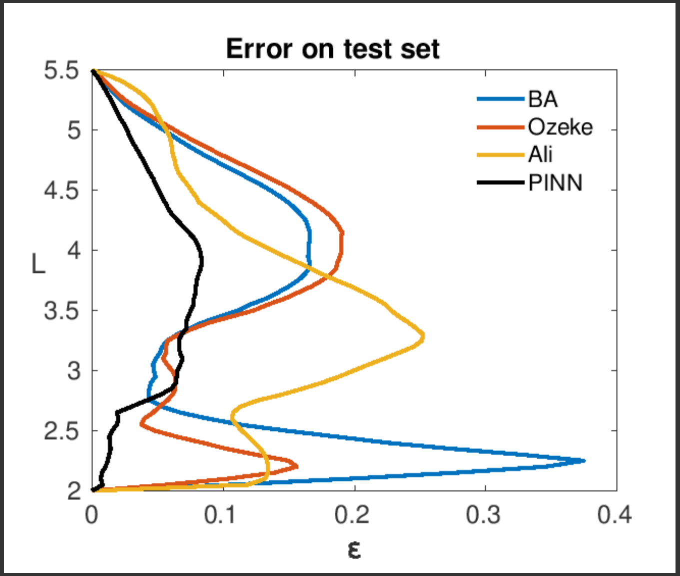
Feature selection

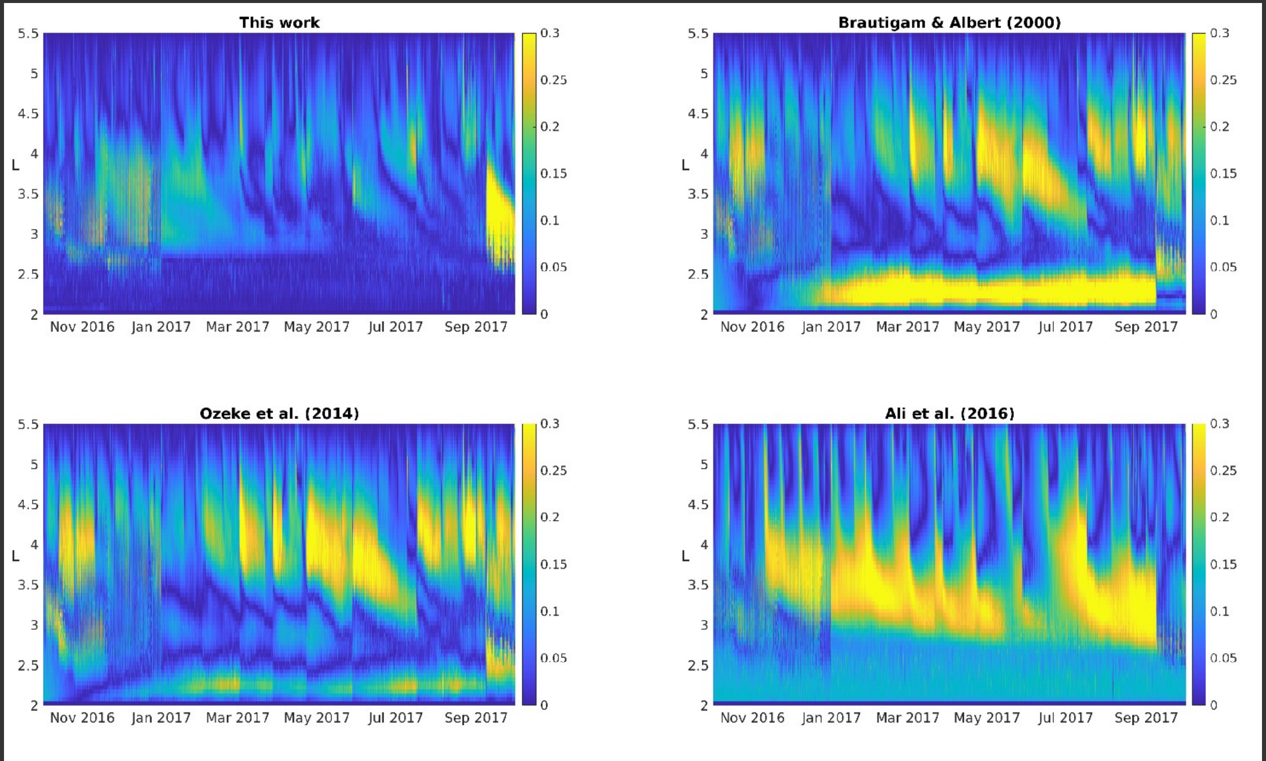


Train a ML model that predicts D_{LL} and C at a given time

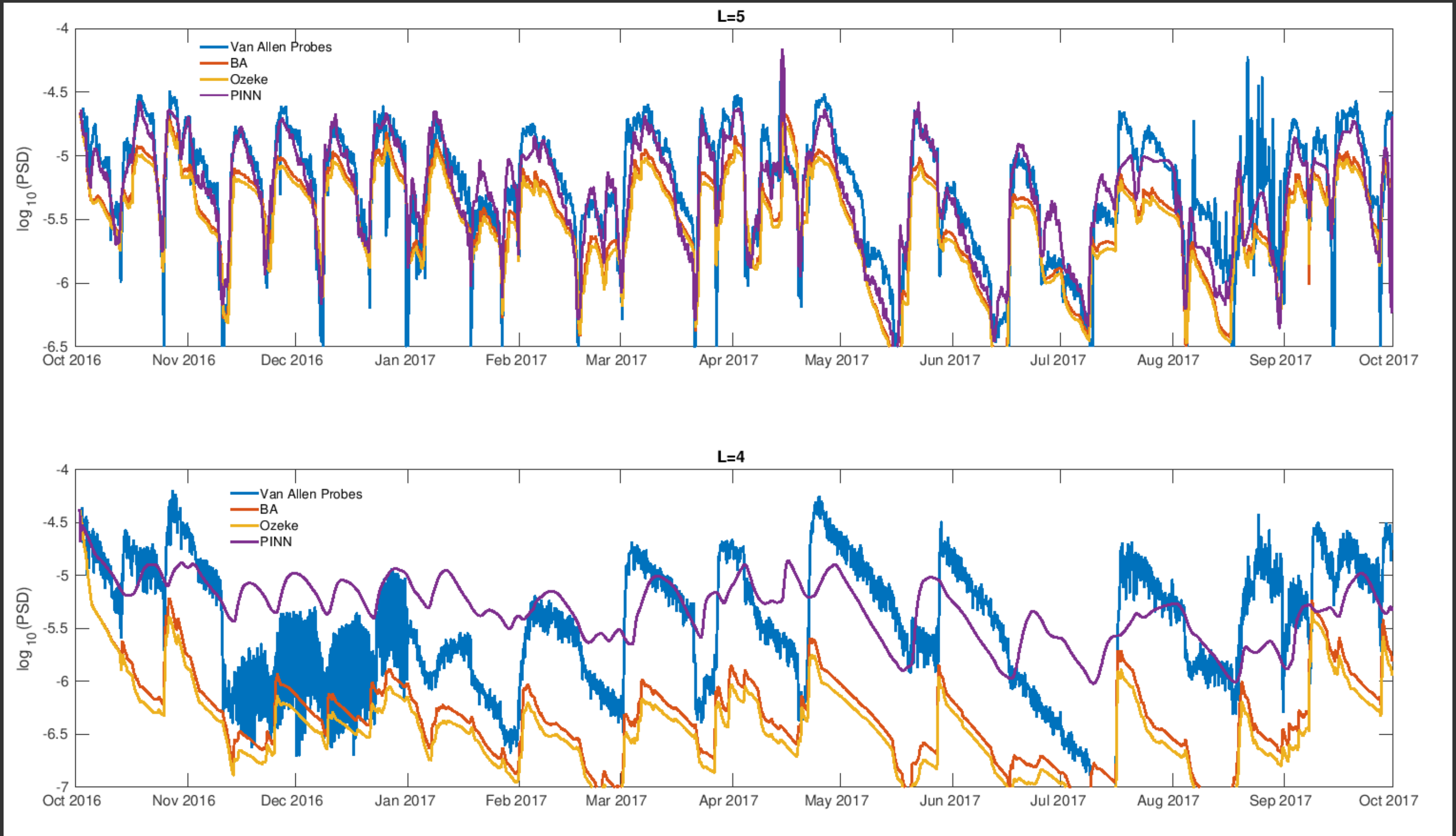
Solve the equation with PINN-discovered and ML-learned coefficients!

Results on test set





Results on test set



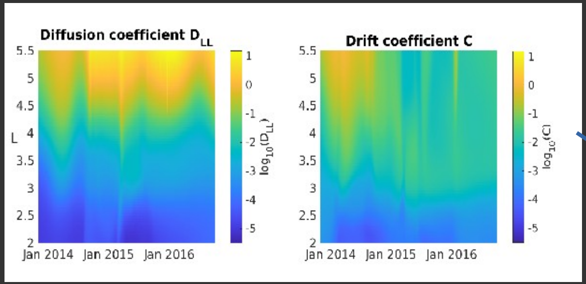
The final frontier: Interpretable AI

Assumption:
The physics obeys FP equation

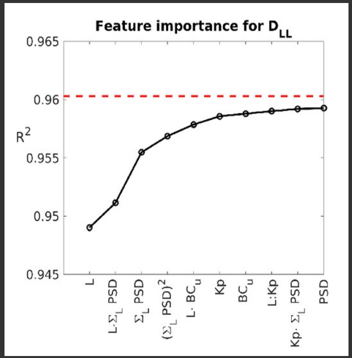
$$\frac{\partial f(L,t)}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial f(L,t)}{\partial L}\right) - \frac{\partial C f(L,t)}{\partial L},$$

PINN

Optimal coefficients

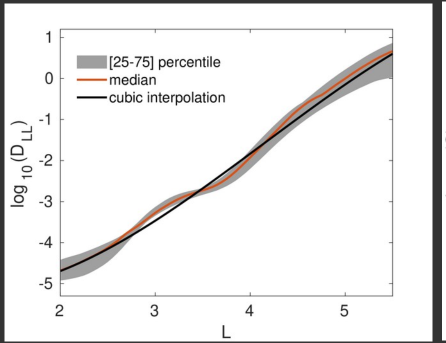


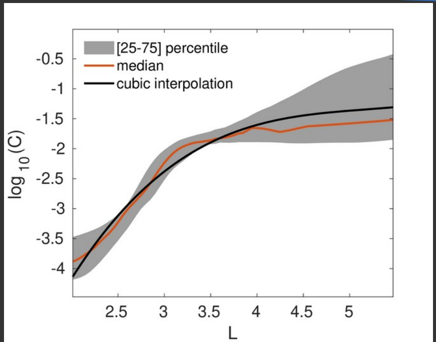
Feature selection



~~Train a ML model that predicts D_{LL} and C at a given time~~

Solve the equation with PINN-discovered and ML-learned coefficients!



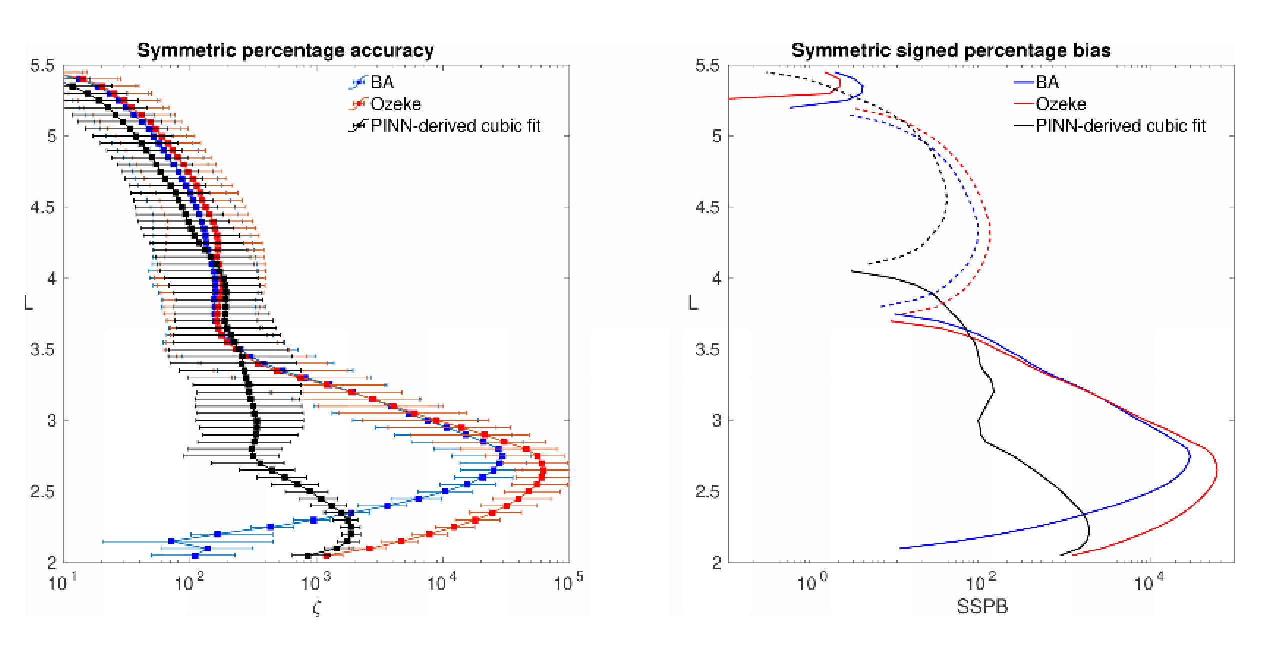


Instead: approximate D_{LL} and C with cubic interpolation

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

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

The final frontier: Interpretable AI



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

D_{LL} and C are functions of L only !!
The FP equation has no free parameters:
Completely determined by BC !!

Diffusion vs Drift

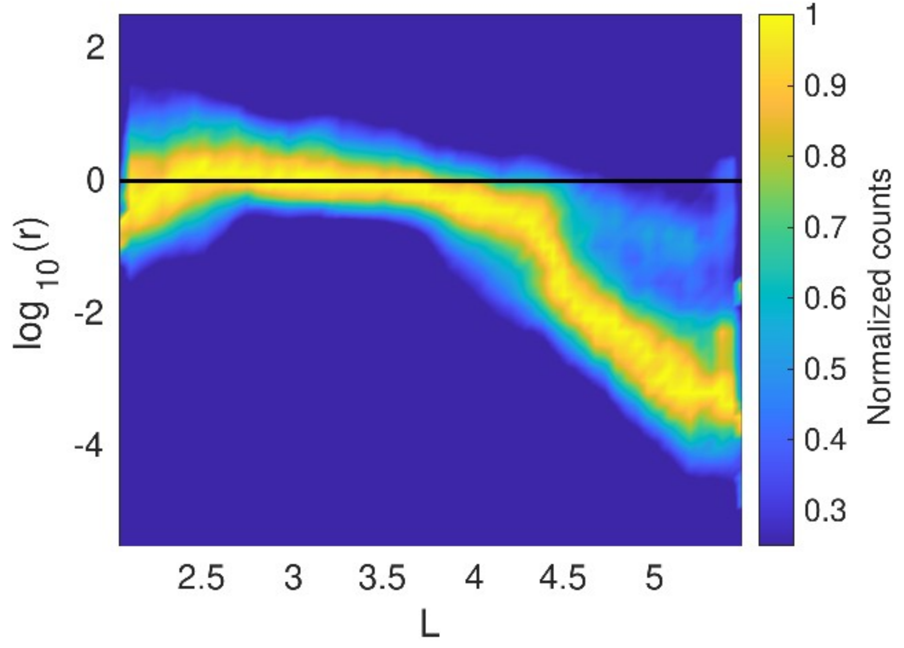


Figure 9. Distribution of r (logarithmic scale) as a function of L. The number of counts is normalized, for each value of L, to its maximum value. The black solid line denotes $r = 1$, that is exact balance between the drift and diffusion terms.

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

Automatic Event Identification

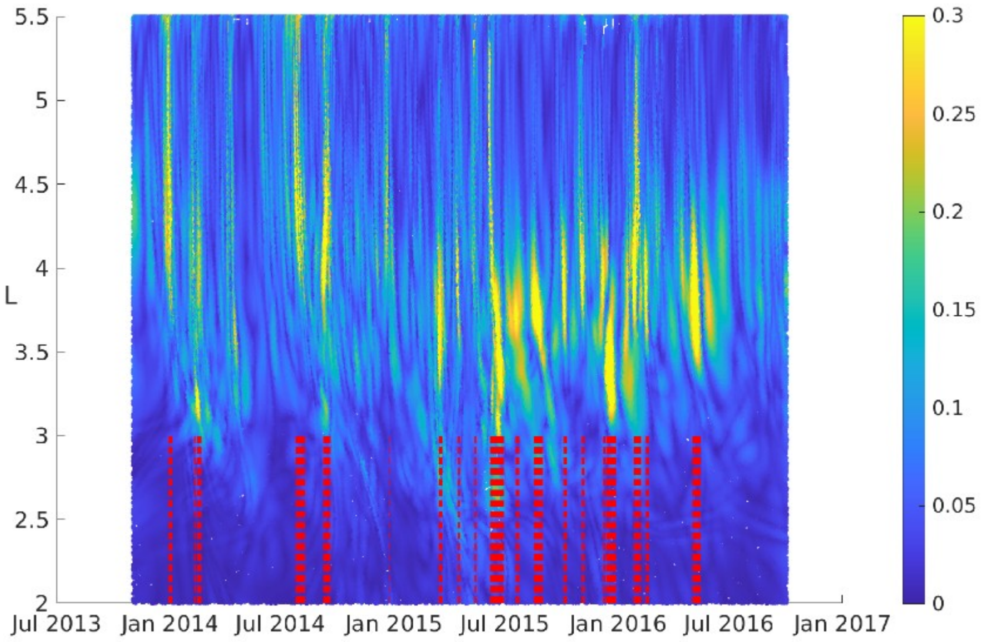


Figure 16. Heat map of the residual of Eq.(4), normalized on its maximum value, over the training set. The red dashed lines denotes time at which the value of the residual is in the 99 percentile of its distribution.

Start time	End time	Previously studied in the literature?
30-Dec-2013	04-Jan-2014	CIR-associated storm (Shen et al., 2017)
10-Feb-2014	10-Feb-2014	
14-Feb-2014	20-Feb-2014	Geomagnetic storm due to multiple interacting ICMEs (Kilpua et al., 2019; Vlasova et al., 2020)
25-Jul-2014	7-Aug-2014	
08-Sep-2014	18-Sep-2014	Dropout event (Ozeke et al., 2017; Alves et al., 2016; Jaynes et al., 2015b; Ma et al., 2020)
24-Dec-2014	24-Dec-2014	
15-Mar-2015	20-Mar-2015	CME-associated storm (Shen et al., 2017; Baker et al., 2016)
15-Apr-2015	17-Apr-2015	
12-May-2015	14-May-2015	CIR-associated storm (Shen et al., 2017)
07-Jun-2015	28-Jun-2015	CIR and CME-associated storms(Shen et al., 2017; Baker et al., 2016); Moderate event(Reeves et al., 2020); Sudden Particle Enhancements at Low L Shells(Turner et al., 2017)
19-Jul-2015	23-Jul-2015	Sudden Particle Enhancements at Low L Shells (Turner et al., 2017)
17-Aug-2015	31-Aug-2015	Moderate event(Reeves et al., 2020)
05-Oct-2015	09-Oct-2015	Moderate event(Reeves et al., 2020)
03-Nov-2015	06-Nov-2015	
08-Dec-2015	11-Dec-2015	
14-Dec-2015	28-Dec-2015	Moderate and strong storms (Boyd et al., 2018b; L.-F. Li et al., 2020; Sotnikov et al., 2019)
27-Jan-2016	07-Feb-2016	Dropout event (Wu et al., 2020)
15-Feb-2016	19-Feb-2016	Moderate event(Reeves et al., 2020); Fast magnetosonic waves(Yu et al., 2021)
01-May-2016	14-May-2016	Moderate event(Reeves et al., 2020; Moya et al., 2017)

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:
  - 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)
  - Re-parametrize the diffusion and drift coefficients with simple, interpretable relationships
  -

Ref: https://doi.org/10.1029/2022JA030377

enrico.camporeale@noaa.gov