3rd URSI AT-AP-RASC, Gran Canaria, 29 May – 3 June 2022

Exploring pitch-angle diffusion during high speed streams with neural networks

J.-F. Ripoll* (1)(2), G. Kluth(1)(2), S. Has(3), A. Fischer(3), M. Mougeot(4), E. Camporeale(5)(6)

(1) CEA, DAM, DIF, F-91297 Arpajon, France; e-mail: jean-francois.ripoll@cea.fr; gilles.kluth@cea.fr
(2) UPS, CEA, LMCE, 91680 Bruyères-le-Châtel, France
(3) LPSM UMR 8001, Université de Paris, France; e-mail: hassothea.math@gmail.com; aurelie.fischer@lpsm.paris
(4) CNRS UMR 9010 Borelli Center, École normale supérieure Paris-Saclay, France; e-mail: mathilde.mougeot@ens-paris-saclay.fr
(5) Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, USA
(6) Space Weather Prediction Center (NOAA), USA; e-mail: enrico.camporeale@noaa.gov

Abstract

This summary paper reports progress on an ongoing development of a neural network (DNN) model of quasi-linear pitch-angle diffusion coefficients of radiation belt electrons caused by whistler-mode hiss waves during high speed streams (HSS). We newly explore pitch-angle diffusion for the HSS of 26 January 2013, discussing the simplicity of performing exploration with a neural network model and comparing how mean diffusion coefficients behave compared with individual ones. The neural network is also newly used to explore the variations of diffusion coefficients at fixed Kp. The model is finally used to predict pitch-angle diffusion during sustained intense HSS yet unobserved.

1. Introduction

Whistler-mode waves in the inner magnetosphere cause electron precipitation in the atmosphere through the physical process of pitch-angle diffusion [e.g. 1]. The computation of pitch-angle diffusion relies on quasi-linear theory and becomes time-consuming as soon as it is performed at high spatial and temporal resolution with the used of event-driven satellite measurements of ambient wave and plasma properties. The effort of integrating detailed measurements of the environment is nevertheless required to capture accurately the variability and complexity of atmospheric electron precipitation [e.g. 2]. In this work, we build a global machine-learning model of event-driven pitch-angle diffusion coefficients for high speed streams conditions based on the data of a variety of storms observed by the NASA Van Allen Probes [3].

2. Machine learning model discrimination method

We first proceed step-by-step by testing 8 nonparametric machine learning methods. We considered methods based on local evaluation (k-nearest neighbors and kernel regression), tree-based methods (regression tree, bagging and random forest), neural networks and function approximations (Radial basis and splines). All are nonparametric so that no assumption is made about the distribution of the data. With these methods, we derive machine learning based models of event-driven diffusion coefficients for the single storm of March 2013 [4] associated to high speed streams (HSS). We define a selection method based on 3 diagnostics that are applied to each method and allows to exhibit selected model properties. The diagnostics are: 1- compute the main global metrics (including mean, median, minimum, maximum, standard deviation and quartiles errors) at various resolutions of the database (here the highest resolution is shown in Table 1), 2- generate violin plots for analyzing the error distribution (as shown in Figure 1(top)), and 3- compute the correlation of each method with the other to enlighten their differences and exhibit the main families (as shown in Figure 1(bottom)).

Table 1. Performances of all the machine learning methods in terms of the absolute error |ei|, reporting the mean error and the standard deviation.

Method	Mean	std
Tree	0.014	0.026
Bag	0.012	0.025
RF	0.012	0.025
KNN	0.010	0.017
KerReg	0.005	0.014
Spline	0.002	0.009
DNN	0.003	0.008

The procedure is repeated on various meshes with different resolution in (energy, pitch angle) and in L-shell. From these diagnostics, we select the best machine learning

method for our problem. Three methods are retained for their accuracy/efficiency: the spline, the radial basis method, and neural networks (DNN), the latter being selected for the second step of the study [5].

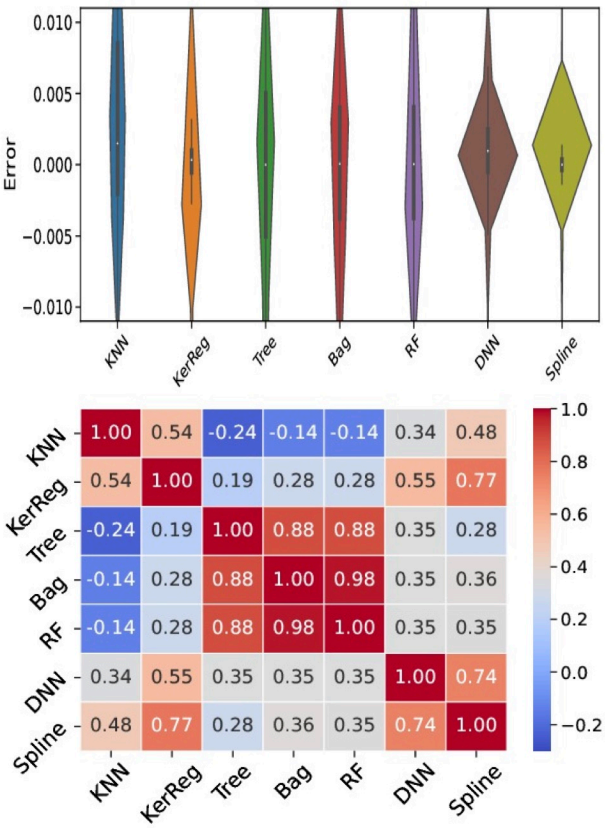

Figure 1. (top) Violinplots of error ei = ui − ûi of all the methods. For each ML-method, the outside envelop is the distribution of error, symmetric for visualization consideration, with a box-whiskers plot inside (median with a white circle, 1st and 3rd quartiles are represented by the border of the box, the whisker showing 95% of the data between percentiles 2.5 and 97.5). (bottom) Correlations of error ei = ui − ûi evaluated on a high resolution (energy, alpha) mesh for 4 L-shells.

3. A neural network model for diffusion coefficients

The retained DNN model is based on the DJINN package and Tensorflow. An original specificity of our DNN model is to optimize the choice of the many hyper-parameters. To do so we use a data-driven method for selecting all the hyperparameters (e.g. [6]). The method uses random forest and a mapping between the obtained trees and the architecture of an ensemble of neural networks. We obtain this way accurate neural networks with only 2 hyperparameters, the depth and the number of trees. We then use event-driven diffusion coefficients which have been previously computed from 32 high speed stream storms observed by the NASA Van Allen Probes at a high temporal and spatial resolution [3]. The raw data represents 1.9E8 values at various L-shell, energy, and pitch angle. Limiting the values to 4 L-shells (L=2,3,4,5) where we test the dataset reduces to 1.8e7 values. We build a first set of averaged diffusion coefficients by considering all the 32 storms, each defined at 9 temporal bins. For a given temporal bin j = 1..9, for a given Kp = 0...6, we average the diffusion coefficients, Dαα(L, E, α), over all the storms. We obtained this way 2.3e6 diffusion coefficients. 2.3e5 values are then used for training and validation and 2.3e5 ones for testing. The dataset contains little noise so that we can train neural networks going deep, with depths of the network going from 6 to 11 hidden layers and an optimum found for a depth of 9. The monotonous variations of the diffusion coefficients as the depths and the data size increase shows that overfitting ends up not to be an issue here as most data are regularized. This way we build for the first time a statistical event-driven diffusion coefficient that is further embedded within the DNN model. We keep both the Kp-index dependence and the storm history in the DNN diffusion coefficients [5]. The double parametrization is chosen to keep both the strength of the storm and its history. In comparison, a Kp-only model is found too inaccurate compared with specific event-driven diffusion coefficients (by 1 to 2 orders of magnitudes depending on the event) (not plotted).

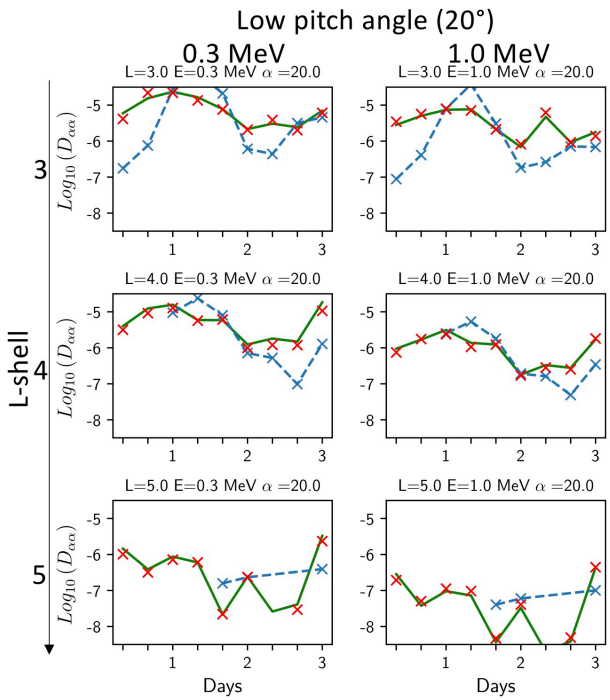

Figure 2. Pitch-angle diffusion coefficients for storm 1 plotted for low pitch angle at 0.3 and 1 MeV. The panels show pitch-angle diffusion coefficient varying with time during 3 days after the storm at different (L, E, α) values, with (blue crosses and lines) the raw data of event-driven coefficient, (red crosses) the averaged data [(on the 32 storms at given (Kp, t, L, E, α))], and (green lines) the DNN model.

4. Pitch-angle diffusion during the HSS of 26 January 2013

In this summary paper, we build on the work presented in [5] by further analyzing results for storm 1 on 26 January 2013 shown in Figure 2 for low pitch angles and Figure 3 for high pitch angles. We see that the average data (in red) and the DNN model (in green) (trained on a subset of the average model) have a very good match for L = 3, 4, 5, signifying the DNN model fulfils its objective. The DNN model reproduces quite perfectly its target, i.e. averaged diffusion coefficients, with rare exception in the Landau resonance region. The DNN mean model is then used to analyze how mean diffusion coefficients behave compared with individual ones for each of the 32 storms. We find a poor performance of mean models (whether they are made of directly averaged data or with DNN) compared with individual events, with mean models recovering the general trend at best. Mean models can easily deviate by 2 orders of magnitude. Loss of accuracy of mean diffusion coefficients comes for the large variance of event-driven diffusion coefficients. This occurs in region of strong gradient of the diffusion coefficients, for instance, at the edge of the first cyclotron resonance in the (E, α) plane (corresponding with pitch angle around ~60° in Figure 3).

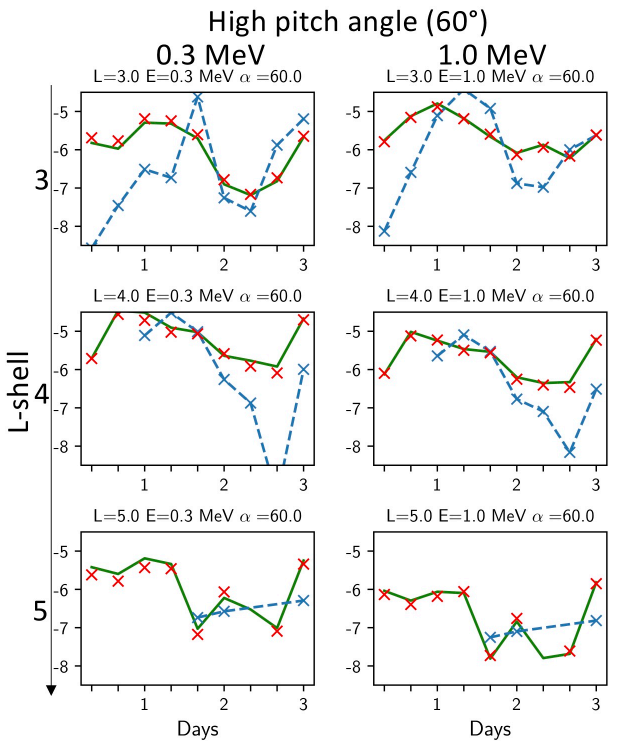

Figure 3. Same as figure 2 for high pitch angle (60°).

5. Exploration of diffusion during HSS

The DNN-based model allows simple and fast data exploration of pitch-angle diffusion among its multiple variables. The strength of the DNN approach is the simplicity of performing comparisons since the model delivers continuous map of the solution with a simple numerical subroutine for a problem with 5 to 6 dimensions here. This is newly illustrated here by model exploration at fixed Kp index, with mild HSS activity (Kp=2) in Figure 4 and high HSS activity (Kp=4) in Figure 5. Pitch-angle diffusion from hiss reduces with Kp increasing. As time evolves, the wave and density properties change from one event to another, even though the Kp index remains the same, showing a multiplicity of solutions for a same Kp. A Kp-only model would retain only a single mean temporal value per L-shell, which cannot be accurate according to the large temporal variability we show here.

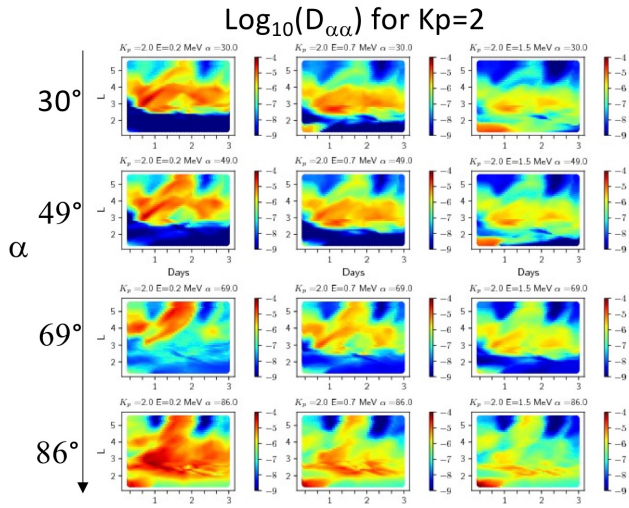

Figure 4. The DNN model of pitch-angle diffusion coefficients (Dαα) (Log10 of s^{-1}) in the (t, L) plane at fixed (α, E) for (top to bottom) α=30, 49, 69, 86° and (from left to right) E=0.3, 0.7, 1.5 MeV for Kp=2.

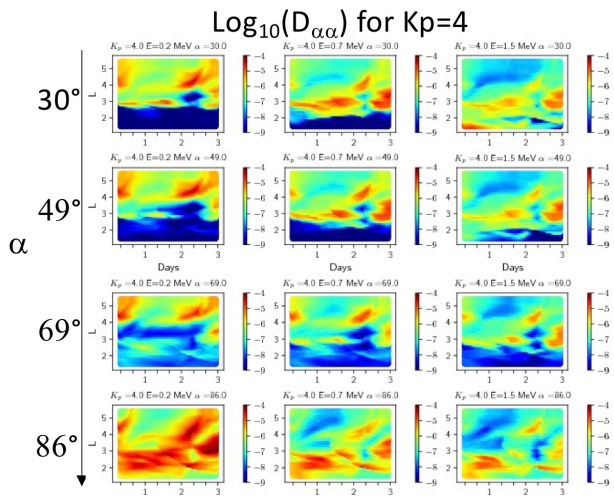

Figure 5. Same as figure 4 for Kp=4.

6. Prediction of pitch-angle diffusion during sustained intense HSS yet unobserved

One advantage of the DNN model is its capacity to extrapolate when values are unknown, producing a prediction. Among the 32 high speed streams, none have a Kp=5 after t=1.3 days (see Figure 8 in [5]), which means

results in Figure 5 are extrapolation for t>1.3 days. The predicted solution is far to be straightforward. We see some of the trend of Figure 3-4 but with very different behaviors. For instance, diffusion is more limited at L~3-4 and α=49-69° than during mild activity. Long lasting intense HSS should thus produce lower diffusion according to these predictions.

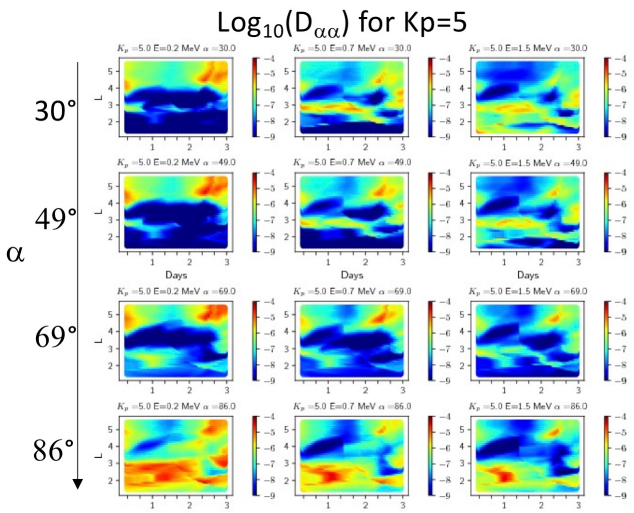

Figure 6. Same as figure 4 for Kp=5. Extrapolation occurs for t>1.3 days.

7. Conclusions

It is difficult to advise the use of a particular machine learning model without prior assumption on a physical phenomenon and its database. One reason is the data used to train the model has a big influence on the model performance, making hard to generalize the capabilities. We define herein a 3-step method based on error statistics, violin plots, and correlation errors, all applied on various meshes, selecting this way the most appropriate method. Three methods are found accurate/efficient: the spline, the radial basis method, and neural networks (DNN). Using neural networks, we achieve building a global DNN mean event-driven pitch-angle diffusion model in which we introduce both a Kp-index and time dependence of the storm history. The neural network model is found accurate at most (L, energy, pitch angle), with loss of accuracy for Landau diffusion. Another strength of the DNN approach is the simplicity of performing comparisons since it delivers continuous maps of the solution with a simple numerical subroutine. Conversely, manipulating directly the discrete points of the data is constraining and source of errors. The neural network is then used to explore the diffusion coefficient database for the high speed stream of 26 January 2013 and analyze how mean diffusion coefficients behave compared with individual ones. The exploration shows poor performance of mean models compared with individual events, with the general trend computed at best, due to the large variability of the wave properties during high speed streams. This variability is further explored looking at the temporal variation of diffusion coefficients even though Kp is fixed. Finally, we explore extrapolated regions of diffusion to predict how pitch-angle diffusion could evolved during sustained intense HSS yet unobserved, finding significant reduction.

8. Perspectives

The method that is proposed here for the DNN mean diffusion can be extended to derive DNN-based median, quartile and standard deviation of the diffusion coefficients. With them, one can perform uncertainty analysis of Fokker Planck simulation and better establish the variability caused by storms and better rank the best possible scenarios for given conditions. We expect this approach to take more importance in the coming years.

References

[1] J.-F. Ripoll, Claudepierre, S. G., Ukhorskiy, A. Y., Colpitts, C., Li, X., Fennell, J., & Crabtree, C., "Particle Dynamics in the Earth's Radiation Belts: Review of Current Research and Open Questions". Journal of Geophysical Research: Space Physics, 125, 2020, https://doi.org/10.1029/2019JA026735

[2] J.-F. Ripoll, Loridan, V., Denton, M. H., Cunningham, G., Reeves, G., Santolik, O., et al.. "Observations and fokker-planck simulations of the l-shell, energy, and pitch angle structure of earth's electron radiation belts during quiet times". Journal of Geophysical Research: Space Physics 124, 1125–1142, 2019. doi:https://doi.org/10.1029/2018JA026111.

[3] Turner, D. L., Kilpua, E. K. J., Hietala, H., Claudepierre, S. G., O'Brien, T. P., Fennell, J. F., et al. (2019). The response of Earth's electron radiation belts to geomagnetic storms: Statistics from the Van Allen Probes era including effects from different storm drivers. Journal of Geophysical Research: Space Physics, 124, 1013-1034. https://doi.org/ 10.1029/2018JA026066

[4] J.-F. Ripoll, J.-F., O. Santolik, G. D. Reeves, W. S. Kurth, M. H. Denton, V. Loridan, S. A. Thaller, C. A. Kletzing, and D. L. Turner (2017), Effects of whistler mode hiss waves in March 2013, J. Geophys. Res. Space Physics, 122, doi:10.1002/2017JA024139.

[5] G. Kluth, J.-F. Ripoll, S. Has, A. Fischer, M. Mougeot, E. Camporeale, "Machine Learning Methods Applied to the Global Modeling of Event-Driven Pitch-Angle Diffusion Coefficients During High-Speed Streams", submitted to Frontiers in Physics, 2021.

[6] G. Kluth, Humbird, K., Spears, B., Peterson, J., Scott, H., Patel, M., et al. "Deep learning for nlte spectral opacities". Physics of plasma 27, 2020.