JGR Space Physics

RESEARCH ARTICLE
10.1029/2019JA027618

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 models based on radial diffusion theory, which is derived from a quasi-linear approach with prescribed inner and outer boundary conditions. The 1D diffusion model requires the knowledge of a diffusion coefficient and an electron loss timescale, which is typically parameterized in terms of various quantities such as the spatial (L) coordinate or a geomagnetic index (e.g., Kp). These terms are typically empirically derived, not directly measurable, and hence are not known precisely, due to the inherent nonlinearity of the process and the variable boundary conditions. In this work, we demonstrate a probabilistic approach by inferring the values of the diffusion and loss term parameters, along with their uncertainty, in a Bayesian framework, where identification is obtained using the Van Allen Probe measurements. Our results show that the probabilistic approach statistically improves the performance of the model, compared to the empirical parameterization employed in the literature.

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

Correspondence to:
R. Sarma,
rakesh@cwi.nl

Citation:
Sarma, R., Chandorkar, M., Zhelavskaya, I., Shprits, Y., Drozdov, A., & Camporeale, E. (2020). Bayesian inference of quasi-linear radial diffusion parameters using Van Allen Probes. *Journal of Geophysical Research: Space Physics*, *125*, e2019JA027618. https://doi.org/10.1029/2019JA027618

Received 22 NOV 2019
Accepted 16 MAR 2020
Accepted article online 7 APR 2020

1. Introduction

The Van Allen radiation belts consist of energetic electrons and protons originating from the solar wind, which are trapped by the Earth's magnetic field. The dynamics of these particles are affected by an interplay of different mechanisms, including local acceleration and losses due to wave-particle interactions and external injections (Reeves et al., 2013; Shprits et al., 2008; 2008; Ukhorskiy & Sitnov, 2012). Understanding and forecasting the radiation belt particles is an essential part of space weather, since these particles can interfere with the satellites orbiting around the Earth and cause data loss. Hence, the estimation of the flux distribution in the space-time domain is a long-standing challenging problem in the space physics community (Reeves et al., 2003; Shprits et al., 2008; 2008).

The standard way of studying radiation belt particles dynamics is through a quasi-linear approach that dates back to the seminal paper of Kennel and Engelmann (1966) and has been routinely applied to magnetospheric particle since the 1970s (Lanzerotti et al., 1970; Schulz & Eviatar, 1969; Schulz & Lanzerotti, 2012).

The quasi-linear approach studies the particles' dynamics based on their quasi-periodic orbits with respect to the field lines of the Earth's magnetic field. Specifically, the cyclotron motion (or gyration) is the circular motion of particles around the field line, the bounce motion is due to the mirror force and along the field lines from one hemisphere to the other, and the drift motion is the orbit in the eastward or westward direction. These orbits are associated to the conservation of adiabatic invariants, respectively, named the first (μ), second (K), and third (L^*) invariant (Roederer, 1970). The evolution of the particle density is then studied as a diffusive process in the adiabatic invariant space, where the effect of resonant wave-particle interactions is described through diffusion coefficients. The third adiabatic invariant is associated with the longest timescale, and it can be violated by wave-particle interactions with ultralow frequency waves. Therefore, as a first approximation, the system can be studied as radially diffusive, assuming the other two invariants to be conserved. In this study, we estimate the phase space density (PSD) based on the case of pure radial diffusion, which has been extensively investigated in the literature.

The radial diffusion formulation involves a parametric representation of the diffusion coefficient and the electron loss timescale, which are generally formulated as varying in L (the spatial coordinate) and Kp (a geomagnetic index used as a proxy for the amplitude of geomagnetic perturbations).

The most extensively employed diffusion rate parameterization is obtained by Brautigam and Albert (2000) based on the October 1990 storm. There have been other developments where the radial diffusion coefficient is evaluated using various approaches. Fei et al. (2006) constructed time dependent diffusion terms from magnetohydrodynamic simulations of September 1998 storm using ultralow frequency electric and magnetic field power spectral density. In Ozeke et al. (2012; 2014), the diffusion coefficient is expressed as a sum of two terms due to azimuthal electric and compressional magnetic fields, and expressions for these two terms are derived. A data-driven approach is employed to determine the radial diffusion coefficients in Su et al. (2015). The importance of the electron loss timescale in the radial diffusion modeling during storm time is demonstrated in Shprits and Thorne (2004), where the interplay between inward radial diffusion and loss terms on the PSD is investigated. Summers et al. (2007) obtained the loss timescale due to the combined effects of chorus, plasmaspheric hiss, and electromagnetic ion cyclotron waves. Shprits et al. (2007) has parameterized the loss timescale due to chorus waves as a function of the geomagnetic index AE. Further extension to include the effect of multiple storms is investigated in Tu et al. (2009) with an internal heating term. In Ali et al. (2016), the magnetic and electric field measurements from the Van Allen Probes are used to compute power spectral densities of both components, which are used with the Fei et al. (2006) formulation to obtain magnetic and electric components of the radial diffusion components.

It is important to emphasize that even in 3D studies of the radiation belt that solve the diffusion equation in the whole adiabatic invariant space, hence taking into account energy and pitch-angle scattering, the radial diffusion coefficient is still often parameterized using the Brautigam & Albert, 2000 formula (see, e.g., Bourdarie & Maget, 2012; Su et al., 2010; Subbotin & Shprits, 2009; Tu et al., 2013; Welling et al., 2012).

However, a single parameterization might not generalize well for different geomagnetic conditions. Moreover, the diffusion problem is significantly influenced by the boundary conditions, hence a deterministic parameterization might not be adequate to represent the uncertainties due to variable particle injections at the boundary. The aim of the current investigation is to obtain a probabilistic representation of the PSD by introducing uncertainties in the parametric representation of coefficients in the underlying partial differential equation. This is one of the first applications of the Bayesian framework approach to parameter estimation of the radial diffusion equation. In the literature, data assimilation with an extended Kalman filter technique is employed for determining the lifetime of electrons by Kondrashov et al. (2007).

In this study, we perform Bayesian parameter identification approach for all the terms defining the 1D diffusion equation. Once the diffusion coefficient and the electron loss time are defined as probability density functions, one can run an ensemble of simulations by sampling different values of the parameters and hence estimate the uncertainty of the output PSD (Camporeale et al., 2016; Camporeale, 2019). We use a data-driven representation of the input parameters, by employing Van Allen Probes data to identify the parameter distribution in a Bayesian setting (Spence et al., 2013).

The manuscript is structured in the following parts: section 2 provides an introduction to the radial diffusion model and the chosen parameterization of coefficients. In section 3, the framework for uncertainty propagation and the Bayesian identification are discussed. The results, discussion, and comparison to the Van Allen measurements and other parameterization in literature are discussed in section 4. Finally, the findings and future research directions are discussed in section 5.

2. Modeling of Radial Diffusion Dynamics

Several radiation belt models (Albert et al., 2009; Li, 2004; Reeves et al., 2012; Subbotin & Shprits, 2009) have been developed to quantify the radial transport. The trapped particles are quantified for given adiabatic invariants (μ, K, L) at time t with the PSD, $f(\mu, K, L, t)$. Under the assumption that the invariants (μ, K) are conserved, the radial diffusion model is a 1D model based on the modified Fokker-Planck equation (Walt, 1970) and is given by

$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left[D_{LL} L^{-2} \frac{\partial f}{\partial L}\right] - \frac{f}{\tau}, \tag{1}$$

where D_{LL} is the diffusion coefficient and τ is the loss timescale, which is essentially a correction term for unaccounted dynamics (such as pitch-angle and energy scattering). Various parameterizations for D_{LL} and τ are available in literature. In particular, in this study, we refer to the seminal work on the statistical diffusion rate of D_{LL} proposed by Brautigam and Albert (2000), which models the geomagnetic storm of 9 October 1990 given by

$$D_{LL}(L,t) = \alpha_D L^{\beta_D} 10^{b_D Kp(t)}, \tag{2}$$

where they use the values $\alpha_D = 4.73 \times 10^{-10}$, $\beta_D = 10.0$, and $b_D = 0.506$. Brautigam and Albert (2000) use a simple estimation of the diffusion coefficients, but their parameterization works surprisingly well even nowadays. However, the equation that is presented here is only one part of the Brautigam and Albert (2000) radial diffusion coefficient, which is the so called *electromagnetic* part. Use of the additional *electrostatic* part leads to incorrect simulation results as was shown by Kim et al. (2011). The choice of the L- and Kp-dependent parameterization for the diffusion coefficient is motivated by the practice in the literature. In the later part of this manuscript, we show that this choice leads to a close agreement between the simulated PSD using equation (1) and that derived from the electron flux measurements from the Van Allen Probes Magnetic Electron Ion Spectrometer (Blake et al., 2013). Further details about the overall radiation belt storm probes mission design, instrumentation, and measurements can be found in Mauk et al. (2013) and Stratton et al. (2013). It has also been reported in earlier investigations (Ali et al., 2016) that D_{LL} exhibited a weak energy dependence in the range of μ between 500 and 5,000 MeV/G. Moreover, in this study we seek to demonstrate the application of Bayesian calibration to the diffusion problem, which can be applied to other choices of D_{LL} parameterization in future investigations.

For the electron lifetime τ, we employ Kp- and L-dependent parameterization based on Gu et al. (2012) (without energy dependence) inside the plasmasphere, while a model based on Ozeke et al. (2014) is used outside the plasmasphere. The plasmapause position L_{pp} is estimated using a recently developed Plasma density in the Inner magnetosphere Neural network-based Empirical (PINE) model (Zhelavskaya et al., 2017). The PINE density model is developed using neural networks and is trained on the electron density data set from the Van Allen Probes Electric and Magnetic Field Instrument Suite and Integrated Science (Kletzing et al., 2013). The model reconstructs the plasmasphere dynamics well (with a cross-correlation of 0.95 on the test set), and its global reconstructions of plasma density are in good agreement with the IMAGE extreme ultraviolet images of global distribution of He_+. The MLT-averaged plasmapause position is calculated using the output of the PINE model by applying a density threshold of 40 cm^{-3} to separate the plasmasphere from the outside of the plasmasphere. The Kp index is obtained from the OMNIWeb database. The parameterization for τ that we employ in equation (1) is given by

$$\begin{aligned} \tau(L,t) &= (\alpha_\tau + \beta_\tau L + b_\tau L^2)/Kp(t) \ \text{ for } \ L \le L_{pp} \\ &= c_\tau / Kp(t) \ \text{ for } \ L > L_{pp}. \end{aligned} \tag{3}$$

We choose a 1-year period from October 2012 to September 2013 for the purpose of analysis. The initial and outer boundary conditions are interpolated from the Van Allen Probes data. As mentioned, equation (1) is obtained for a constant value of (μ, K). In this study, we compare all the results with Van Allen measurements for $\mu = 700$ MeV/G and $K = 0.0019$ $G^{0.5}$·Re. In order to estimate the accuracy of our predictions, we use the relative error as a metric given by

$$\epsilon = \frac{|f_{va} - f|}{f_{va}}, \tag{4}$$

where f_{va} is the PSD value obtained from the Van Allen Probes and f is the PSD estimate from equation (1). The absolute value of the discrepancy is used here, since we are interested in estimating the overall performance of the solver throughout the domain, which will be obtained by integrating this error across the domain. Now, as defined in section 1, we are interested in obtaining an informed estimate on the parameters defining the coefficients D_{LL} and τ in equation (1). In the next section, we introduce the Bayesian framework which is used to identify these parameters.

Table 1
Upper and Lower Bounds of Prior Defined in Equation (6) for Seven Parameters in $\mathbf{\Lambda}$

Parameter	Lower bound	Upper bound
α_D	0.0	1.4×10^{-7}
β_D	1.0	45.0
b_D	0.0	13.0
α_τ/Kp	$0.0/Kp$	$26.0/Kp$
β_τ	0.0	5.0
b_τ	0.0	5.0
c_τ	0.0	30.0

3. Bayesian Framework for Identification

The Bayesian approach to the calibration of computer models was introduced in Kennedy and O'Hagan (2001), where the term calibration refers to adjusting the free parameters in order for the model output to fit the observations. In our case, the forward model is represented by equation (1) that is solved numerically with a standard finite-difference scheme on a uniform grid in (L, t). Because solving this equation numerically is relatively fast, we opt for a standard Markov Chain Monte Carlo (MCMC, Brooks et al., 2011) procedure to explore the space of unknown free parameters in equations (2) and (3) that are collected in a multidimensional vector $\mathbf{\Lambda}$ defined as

$$\Lambda := (\alpha_D, \beta_D, b_D, \alpha_\tau, \beta_\tau, b_\tau, c_\tau).$$

In other words, $\mathbf{\Lambda}$ is the set of uncertain parameters that are to be identified from this investigation. The ground truth for the PSD is taken from Van Allen Probes measurements, and we derive a Bayesian model trained over a dataset of 30 days from 01 to 30 October 2012. It is to be noted that the choice of this time interval is arbitrary and the objective is to demonstrate that the method is generalized for time periods on which it is not trained, hence the model will be tested for the rest of the year (November 2012 to September 2013). The vector of parameters $\mathbf{\Lambda}$ is treated as a random variable, meaning that it is associated to an (unknown) probability density. The scope of the Bayesian inference is to estimate the probability of $\mathbf{\Lambda}$, given the PSD observations, that we denote with $\mathbf{f}^+$. Hence, we can use the classical Bayes' rule (Gelman et al., 2004):

$$\mathbb{P}(\Lambda|\mathbf{f}^+) \propto \mathbb{P}(\mathbf{f}^+|\Lambda) \cdot \mathbb{P}_0(\Lambda), \tag{5}$$

where $\mathbb{P}(\mathbf{f}^+|\Lambda)$ is the *likelihood* that defines the discrepancy between the model estimate and the Van Allen Probes measurements, for a given realization of $\mathbf{\Lambda}$. The term $\mathbb{P}_0(\Lambda)$ is the *prior* distribution of $\mathbf{\Lambda}$, which encodes all prior physical information one might have about the parameters. $\mathbb{P}(\Lambda|\mathbf{f}^+)$ is called the *posterior* distribution, which in general cannot be expressed in closed form but can be sampled through a Monte Carlo procedure (see below). Finally, once a sufficient number of samples from the posterior distribution has been collected, an ensemble simulation can be obtained by propagating each realization of $\mathbf{\Lambda}$ through equation (1) to obtain the posterior predictive distribution of the PSD.

3.1. Prior, Likelihood, and Posterior

As a first step, the priors on the parameter set $\mathbf{\Lambda}$ are defined. In the present investigation, we assume uniform priors given by

$$\Lambda \sim \mathcal{U}(\Lambda_{min}, \Lambda_{max}), \tag{6}$$

where $\mathbf{\Lambda}_{min}$ and $\mathbf{\Lambda}_{max}$ are chosen such that a wide domain is defined. The bound for each parameter of $\mathbf{\Lambda}$ is shown in Table 1. Existing parameter values identified in literature (Brautigam & Albert, 2000; Ozeke et al., 2014) are used as a reference, such that parameterizations obtained in these investigations belong to the set of the defined bounds.

We employ a Gaussian likelihood, by introducing an additive random variable η, specified as an unbiased normal distribution with a covariance matrix Σ. Under the assumption that there are no modeling errors, the statistical model can then be written as

$$\begin{aligned} \mathbf{f}^+ &= \mathbf{f}(\Lambda) + \eta \\ \eta &\sim \mathcal{N}(0, \Sigma), \end{aligned} \tag{7}$$

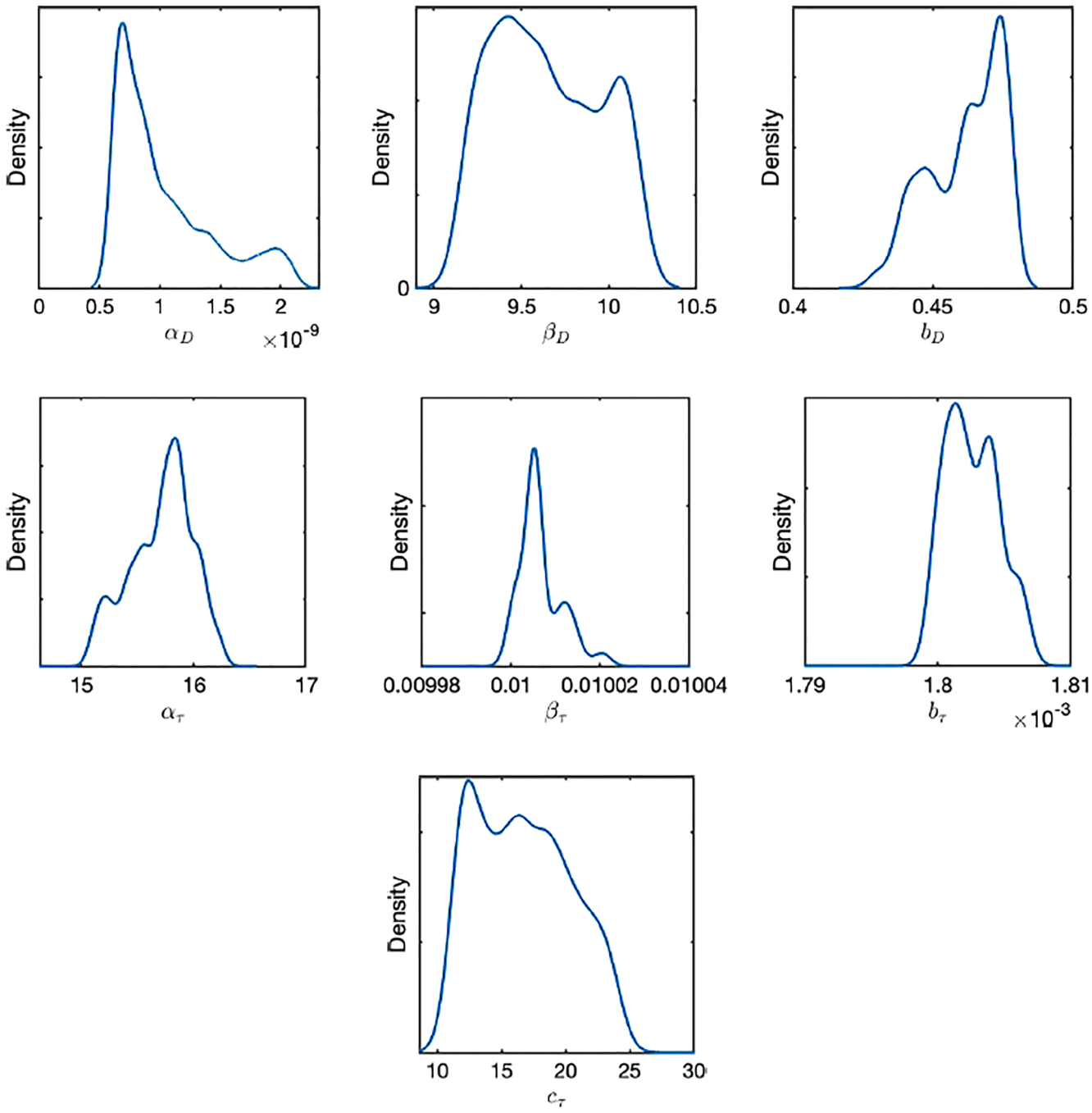

Figure 1. Posterior distributions of identified parameters $\mathbf{\Lambda}$ of D_{LL} and τ.

where $\mathbf{f}(\mathbf{\Lambda})$ is the output of the forward model (1), which gives the likelihood as

$$\mathbb{P}(\mathbf{f}^+|\Lambda) := \mathbb{P}_\eta(\mathbf{f}^+ - \mathbf{f}(\Lambda)) = |\Sigma|^{-n/2} \exp\left[-\frac{1}{2}\left(\mathbf{f}^+ - \mathbf{f}(\Lambda)\right)^T \Sigma^{-1} \left(\mathbf{f}^+ - \mathbf{f}(\Lambda)\right)\right], \tag{8}$$

where $|\Sigma|$ is the determinant of the covariance matrix and n is the number of Van Allen probe measurements used to train the model.

For sampling the posterior, we use the Metropolis-Hastings MCMC algorithm (Hastings, 1970), with a reversible Markov random walk. The algorithm performs a random walk by sampling from a proposal distribution with a prescribed variance. Each new sample distribution is conditional only on the current sample, hence the generated sequence of samples resembles a Markov chain. Once a sample is chosen, a ratio of

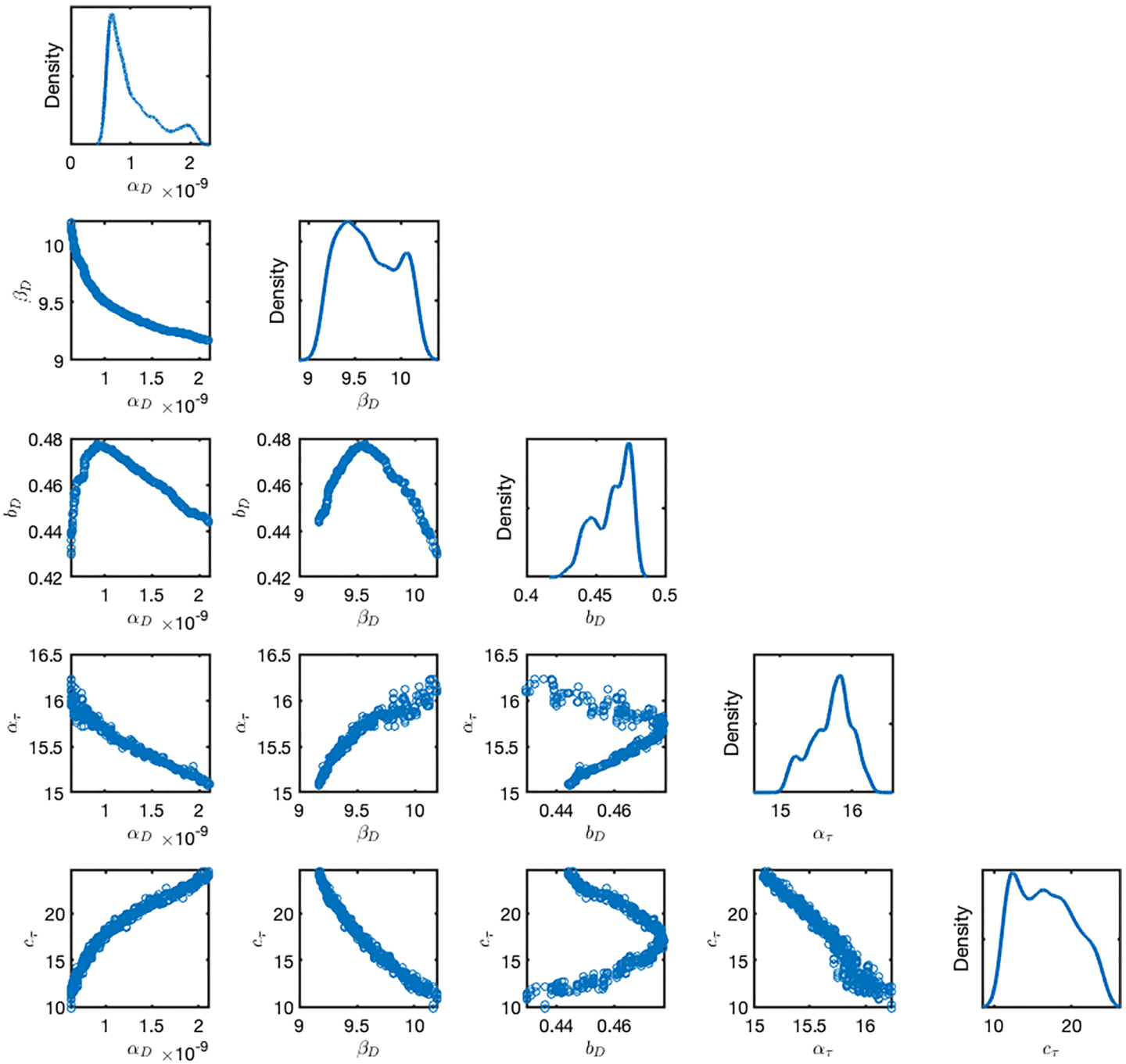

Figure 2. Identified 1D posterior densities and 2D marginal distributions for five parameters defining the diffusion and loss terms in equation (1).

the densities for the two consecutive samples is computed, which informs the iterator if the jump results in increase/decrease of the posterior density. If the jump results in a probability higher than an acceptance limit, the sample is accepted with that probability or it is rejected. The proposals are selected from a normal distribution, with tuned variance and a total of 20,000 samples, where the initial 1,000 samples are discarded as *burn-in* period. And a thinning factor of 2 (Gelman et al., 2004) is applied for the remaining samples to improve independence. It is to be noted here that choice for the burn-in period and thinning factor is problem specific, and it is decided based on the convergence of the parameters. The standard checks of convergence and autocorrelation for MCMC were applied (Chib & Greenberg, 1995; Hastings, 1970).

Table 2
Comparison of Parameterization Based on the Maximum a Posteriori (MAP) Estimate of the Identified Posterior to 1D Model in Drozdov et al. (2017)

Parameter	MAP estimate	Reference
α_D	7.06×10^{-10}	4.73×10^{-10}
β_D	9.43	10.0
b_D	0.47	0.51
α_τ/Kp	$15.84/Kp$	10.0
β_τ	0.01	-
b_τ	0.0018	-
c_τ	12.22	6.0

4. Results and Discussion

In this section, we show the application of the Bayesian inference for identifying the parameters in equation (1), which are updated with the Van Allen Probes measurements. The 7-parameter uncertain space given by $\mathbf{\Lambda}$ is propagated with the forward solver and then the Bayesian framework is applied.

4.1. Posterior Distribution of Diffusion and Lifetime Parameters

Figure 1 shows the posterior distributions of the identified parameters $\mathbf{\Lambda}$ of D_{LL} and τ. As mentioned, uniform priors are specified with wide domains, where, for example, the prior of $\beta_D \sim \mathcal{U}(1, 45)$. It is observed that all the identified posterior distributions have low variance, except for the parameter c_τ, whose uncertainty is still high after identification, which shows that the data is not informative enough to infer c_τ. It is also interesting to note that c_τ in equation (3) is employed above the plasmapause location L_{pp}, which implies that this term is mostly employed at higher values of L, where the flux from the boundary dominates the density predictions. As a result, the inferred values of c_τ strongly depend on the influx of particles at the outer boundary, which is highly uncertain and a deterministic representation is not possible.

Furthermore, parameters β_τ and b_τ have values close to zero, which are coefficients for L and L^2 dependence terms of τ in equation (3), respectively. This shows the variation of τ with respect to L is minimal. It should also be mentioned here that other investigations of higher order variation of τ with L also showed negligible dependence.

In Figure 2, 2D marginal distributions of each of the parameter pairs for five of the identified parameters in $\mathbf{\Lambda}$ are plotted along with the 1D probability densities. The parameters β_τ and b_τ are not plotted here since their identified samples are close to zero. It is interesting to observe that clear correlations are identified between many parameter pairs. Parameters α_D and β_D have a clear negative correlation with an asymptotic trend. Furthermore, parameter pairs α_D and α_τ, β_D and c_τ, α_τ and c_τ have negative correlation, while α_D and c_τ seem to have a positive correlation. Other parameter pairs do not seem to show a clearly visible correlation behavior.

For the identified posteriors, we determine the maximum a posteriori (MAP) estimate of the probability distribution, which is the parameter value corresponding to the point of maximum probability density. Table 2 compares the MAP estimate of the posteriors to the parameterization employed in the 1D model of Drozdov et al. (2017), where two parameterizations for the 1D diffusion problem are presented. In this study, we compare our estimates to the 1D model based on Brautigam and Albert (2000) diffusion parameterization in Drozdov et al. (2017). Analytical Kp-dependent model based on Shprits et al. (2005) was employed outside the plasmasphere, where $\tau = 3/Kp$ days is used. In order to take into account the effect of the variable outer boundary, Shprits et al. (2006) used $\tau = 5/Kp$ as some of the loss was provided by the outward diffusion and longer lifetimes resulted in a better comparison with observations. Since we have a variable boundary in this investigation, we compare our estimates to a parameterization with a longer lifetime similar to Drozdov et al. (2017) and Ozeke et al. (2014), given by $\tau = 6/Kp$ days. It is to be noted that for the subsequent sections of this manuscript, we use the same choice of parameterization as reference.

Figure 3 compares the probabilistic D_{LL} obtained with the identified posteriors shown in Figure 1 to the D_{LL} employed in Brautigam and Albert (2000). At low values of Kp, the two predictions are close to each other,

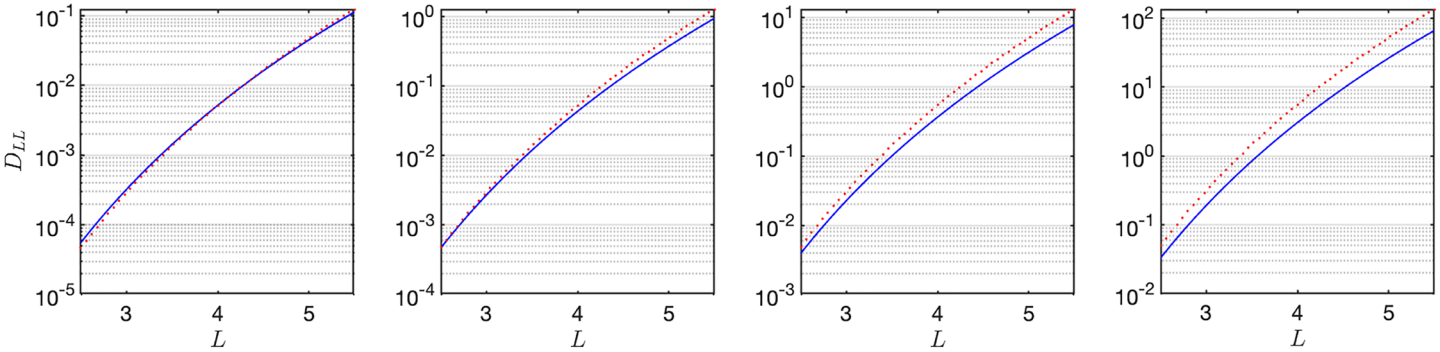

Figure 3. Probabilistic representation of D_{LL} (in log scale) with the identified posteriors, with $Kp = 2$, 4, 6, and 8 from left to right. The red dots represent D_{LL} obtained with Brautigam and Albert (2000) parameterization, while the blue solid lines represent the mean of the probabilistic D_{LL} obtained from the posteriors.

while the discrepancy between them progressively increases as Kp is increased (note the vertical logarithmic scale). With Brautigam and Albert (2000) parameterization, D_{LL} has higher values at higher L with increase in Kp compared to the probabilistic treatment. This could be related to the Kp-dependent τ parameterization, both above and below the plasmapause location.

It is observed that the MAP estimates for the parameters of D_{LL}, namely, α_D, β_D, and b_D, have values close to the reference parameterizations. For β_τ and b_τ, reference values do not exist and they both are close to zero. In equation (3), α_τ is Kp dependent, while in the reference it has a constant value of 10.0 below and $6.0/Kp$ above the plasmapause location. With the current choice of parameterization, this term would be significantly smaller for high Kp while being comparable to the reference value for low Kp. Furthermore, the term c_τ also significantly varies with respect to the reference value. As discussed already, c_τ has high variance, and the MAP estimate is higher than the reference. Hence, it is observed that in terms of the MAP comparison, the lifetime estimates are giving longer time scales than that estimated in literature. This is clearly due to omission of local acceleration in the 1D diffusion model. Similarly, the 1D model in Drozdov et al. (2017) underestimated the observations. Hence, the Bayesian parameter estimation is trying to compensate for the missing physical processes.

4.2. Probabilistic Representation of PSD

Here we investigate the effect of the predicted uncertainties of the diffusion parameters on the evolution of the PSD. In practice, because the posterior distributions are derived by $\sim 10,000$ collected samples, one can solve the diffusion equation (1) for every realization of the parameters, hence generating an ensemble of runs from which to derive the uncertainty of the PSD.

4.2.1. Probabilistic Variation of Density

In order to visualize the probabilistic field of density in the (L, t) domain, we plot the PSD at various time instances. Figure 4 shows the PSD variation at multiple time snapshots, advancing in time from Figures 4a to 4j. The solid line represents the mean of the posterior predictions, while the shaded area shows the 1σ confidence interval. At the beginning of the simulation near the initial condition (Figure 4a), it can be observed that the uncertainty is low. Also there is a good agreement with respect to the Van Allen Probe measurements (red dots), which is expected since the problem is initialized with the data. The low uncertainty is also expected since the identified posteriors have low variance.

The solution is further time marched and the uncertainty progressively increases (Figures 4b–4e). It can be observed that the uncertainty near the upper boundary is smaller, while it is higher around the lower boundary at all time instances. This is due to the influence of the upper boundary condition, which is interpolated from the Van Allen Probe data. It is interesting to observe the influence of the injection of particles from the outer boundary. In Figure 4f, the uncertainty around $L = 4$ is observable, while after the influx of particles in Figure 4g, the uncertainty reduces significantly. Also the Van Allen Probe data have significant noise and discontinuities in Figure 4g, since higher density gradients diminishes the smoothness of interpolation. The uncertainty again starts to increase in Figure 4h, and a similar effect with injection of particles can be observed in Figures 4i–4j. The red dots indicate the interpolated value of the Van Allen Probes data at a given time. Due to the effect of interpolation, it is not expected that the simulation outputs and the red dots would coincide perfectly. Overall the model performs well in terms of agreement with the Van Allen

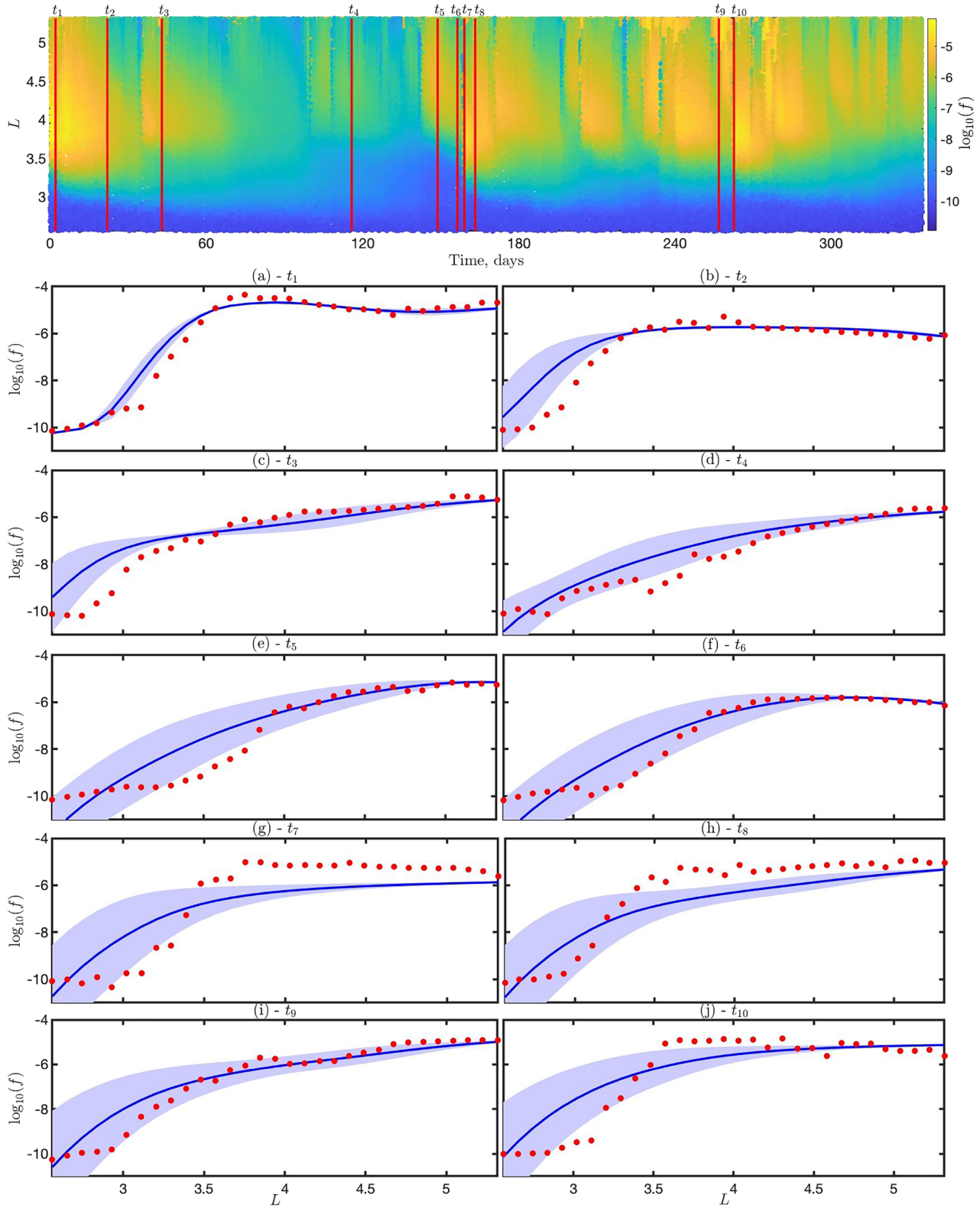

Figure 4. Snapshots of phase space density at $\mu = 700$ MeV/G and $K = 0.0019$ G$^{0.5}$·Re for different time instances. Top panel shows the scatter plot of Van Allen Probes measurements with red lines indicating the time instances at which the probabilistic response is plotted, advancing in time from (a) to (j). Each subfigure from (a) to (j) shows the posterior-propagated phase space density with mean (solid blue line), 1σ confidence intervals (blue shade), and the Van Allen measurements (red dots) corresponding to the time instance t_1 to t_{10} in the top panel.

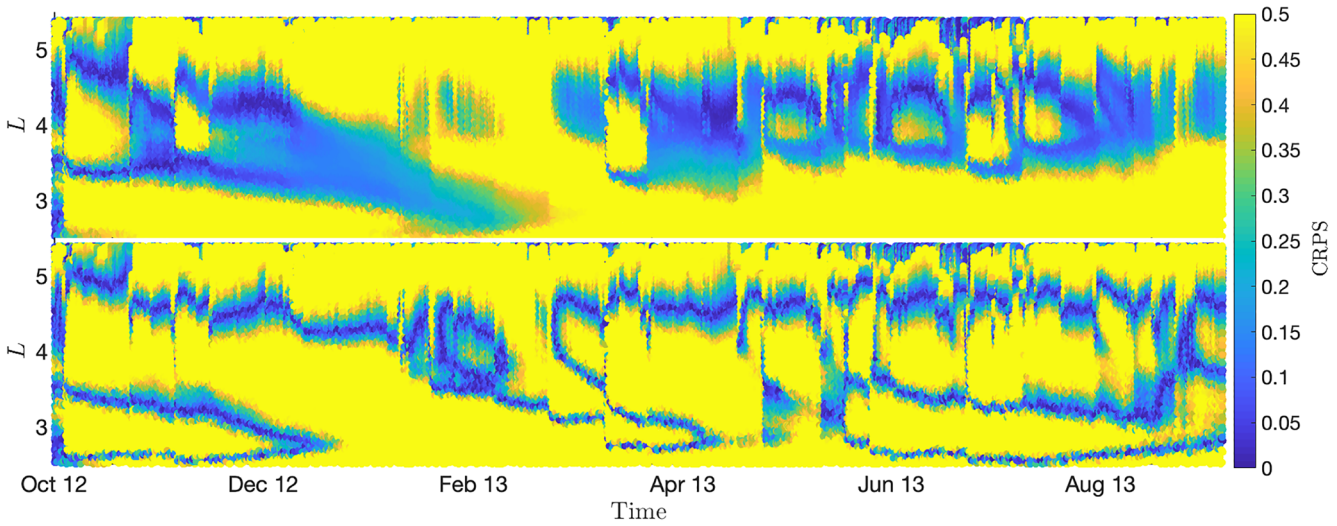

Figure 5. Scatter plot of continuous rank probability score (CRPS) for the time period from October 2012 to September 2013 from Van Allen Probes. The top panel correspond to the probabilistic predictions of phase space density by equation (1) with the identified posteriors, while the bottom panel is obtained from the deterministic predictions of equation (1) with the reference parameterization (Drozdov et al., 2017).

Probes measurements, often predicting within one standard deviation. Furthermore, the lack of agreement at lower values of L is accounted by the confidence intervals. It is observed that the model discrepancy is higher when there is sudden influx of particles, which cannot be properly accounted for in a diffusive process. However, with our probabilistic approach, we are able to obtain overlap with respect to the data at most values of L.

4.2.2. Continuous Rank Probability Score

In order to measure the performance of the probability forecasts, we employ the continuous rank probability score (CRPS), which is a metric to evaluate the quadratic discrepancy between the cumulative distribution function F of the forecasts and the empirical cumulative distribution function of the data or observations F^d, which is given by

$$CRPS = \int_{-\infty}^{+\infty} \left(F(f) - F^d(f)\right)^2 df, \tag{9}$$

where F^d is a Heaviside function for a scalar observation. Thus, CRPS is computed by integrating the area between the two distribution functions. If the computed area is low, the forecasts are close to the observations. Also, CRPS is a convenient measure to compare the performance of deterministic and probabilistic forecasts (Camporeale et al., 2019). If the forecast is deterministic, CRPS reduces to a mean absolute error. Because we have a probability density of the PSD for any point of the domain (L, t), we can numerically calculate the CRPS over the whole domain. Figure 5 compares the CRPS obtained from the probabilistic predictions of PSD with the identified posteriors shown in Figure 1 and the deterministic predictions with the parameterization referenced in Table 2. With respect to assessing the quality of the predictions, values closer to zero show that the predictor is performing better. A qualitative comparison shows that the probabilistic forecast performs better at most of the locations in the (L, t) domain. Figure 6 shows the comparison of the density of CRPS in the whole domain with the probabilistic and deterministic approaches. The mean of the CRPS is 0.429 and 0.499, while the 1 standard deviation is 0.4 and 0.46 for the probabilistic and deterministic solver, respectively. From Figure 6, it can be seen that the probabilistic approach returns a stronger peak meaning it performs better at most sample locations in the domain. It can thus be concluded that the probabilistic treatment improves the performance of the solver.

4.3. MAP Estimates of Phase Space Density

The advantage of the proposed method clearly stays in the ability of deriving a probabilistic estimate of the PSD and the consequent ability of identifying cases of large uncertainty. However nothing precludes to use the information on the posterior distribution of the parameters in a deterministic way, by running a simulation corresponding to the MAP estimate of the parameters. In this section, we compare the PSD at

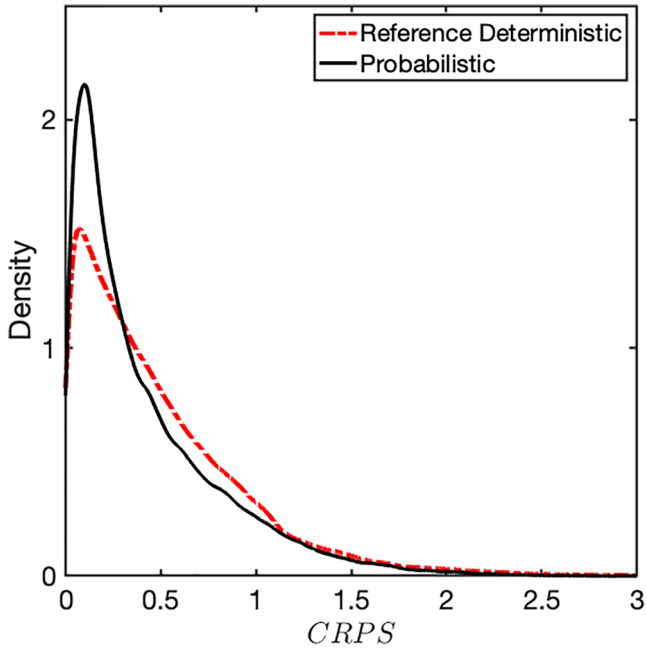

Figure 6. Density plot of continuous rank probability score (CRPS) samples shown in Figure 5. The black solid line correspond to the probabilistic predictions of phase space density by equation (1), while the red dotted line correspond to predictions by the reference parameterization (Drozdov et al., 2017).

$\mu = 700$ MeV/G and $K = 0.0019$ G$^{0.5}$·Re, predicted by the MAP estimate of the identified posteriors of $\mathbf{\Lambda}$ to the Van Allen Probe measurements. As a reference we again employ the 1D model comparison provided in Drozdov et al. (2017) (see Table 2). Additionally, we compare the predictions to other parameterizations existing in literature. We define D_{LL} based on parameterizations obtained in Ozeke et al. (2014) and Ali et al. (2016). In Ozeke et al. (2014), D_{LL} is defined as the sum of azimuthal electric field D_{LL}^{E} and compressional magnetic field D_{LL}^{M} given by

$$\begin{aligned} D_{LL}^{E} &= L^{6} 2.6 \cdot 10^{-8} 10^{(0.217L + 0.461Kp)}, \\ D_{LL}^{M} &= L^{8} 6.62 \cdot 10^{-13} 10^{(-0.0327L^{2} + 0.625L - 0.0108Kp^{2} + 0.499Kp)}. \end{aligned} \tag{10}$$

In Ali et al. (2016), the drift-averaged power spectral densities are used to obtain the electric and magnetic components of the radial diffusion coefficient, and a genetic algorithm is used to derive a simple model for each component in the least squares sense, which is given by

$$\begin{aligned} D_{LL}^{E} &= \exp(-16.951 + 0.181 \cdot Kp \cdot L + 1.982 \cdot L), \\ D_{LL}^{M} &= \exp(-16.253 + 0.224 \cdot Kp \cdot L + L). \end{aligned} \tag{11}$$

Figure 7 shows the PSD in the form of scatter plots and compares five different realizations. It can be observed that the forward model estimations have good qualitative agreement with the Van Allen measurements. The model predictions by equation (1) with the MAP-based and the reference parameterizations are in close agreement. This is expected since the MAP estimate of the posteriors, in particular the diffusion parameters, match closely with the reference values, which can be seen in Table 2. The PSD estimates from the diffusion model based on equation (10) (Ozeke et al., 2014) are also close to the MAP-based model predictions. For the diffusion model based on equation (11) (Ali et al., 2016), the PSD predictions at higher values of L are in qualitative agreement; however, there is higher discrepancy at lower L values. It has to be mentioned here that the model (11) was calibrated in the range $3.0 \leq L \leq 5.5$ and $0 \leq Kp \leq 5$. Hence, in this case we are performing extrapolation, which may lead to lower performance of the model.

We also compare the relative error given by equation (4) in estimation of PSD with scatter plots shown in Figure 8. As expected, predictions from the MAP-based and reference parameterization match closely. The MAP-based model performs better at lower values of L. Similarly, the model from Ozeke et al. (2014)

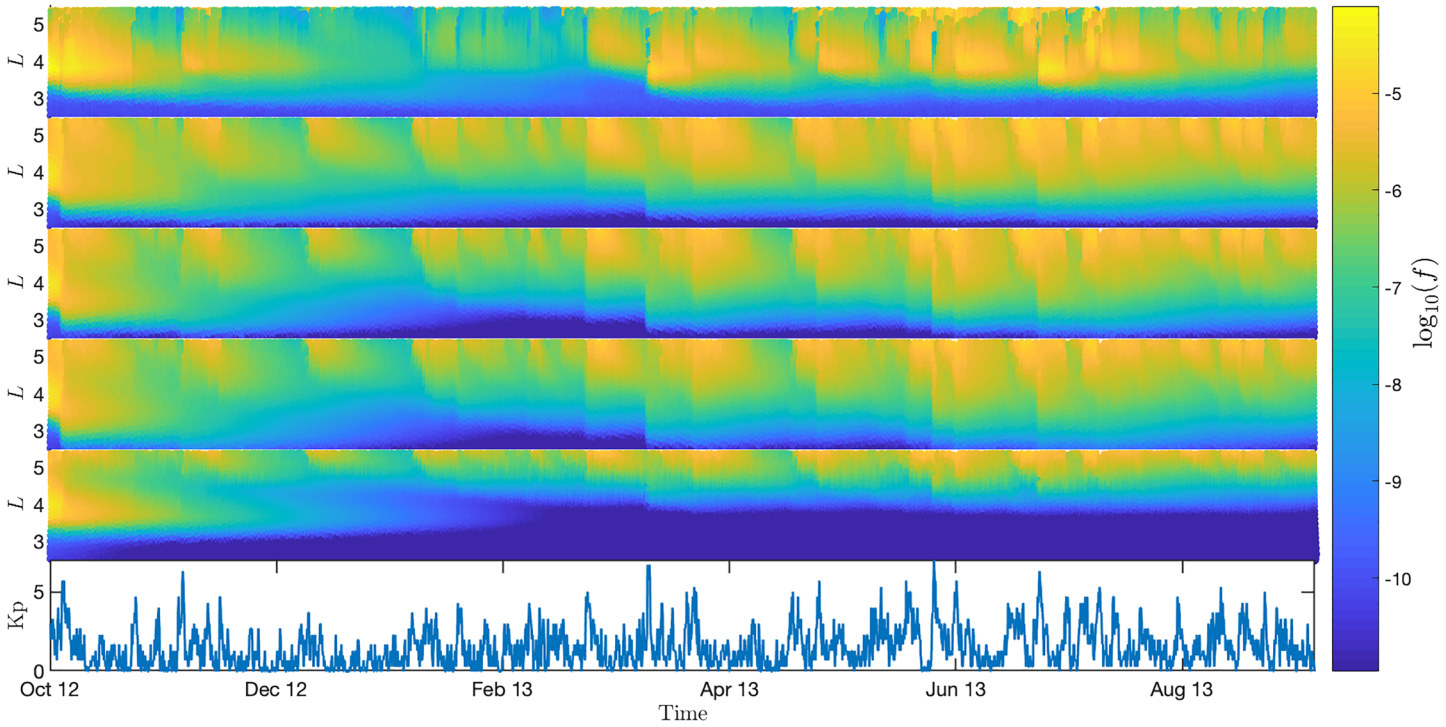

Figure 7. Scatter plot of phase space density (in logarithms of base 10) at $\mu = 700$ MeV/G and $K = 0.0019$ $G^{0.5}$·Re in $L - t$ domain for the time period from October 2012 to September 2013 obtained from Van Allen Probes measurements (top panel), model predictions by equation (1) and parameterized by the maximum a posteriori estimate of posteriors (second panel), reference (Drozdov et al., 2017) (third panel), equation (10) from Ozeke et al. (2014) (fourth panel), equation (11) from Ali et al. (2016) (fifth panel), and *Kp* index from OMNIWeb database (bottom panel).

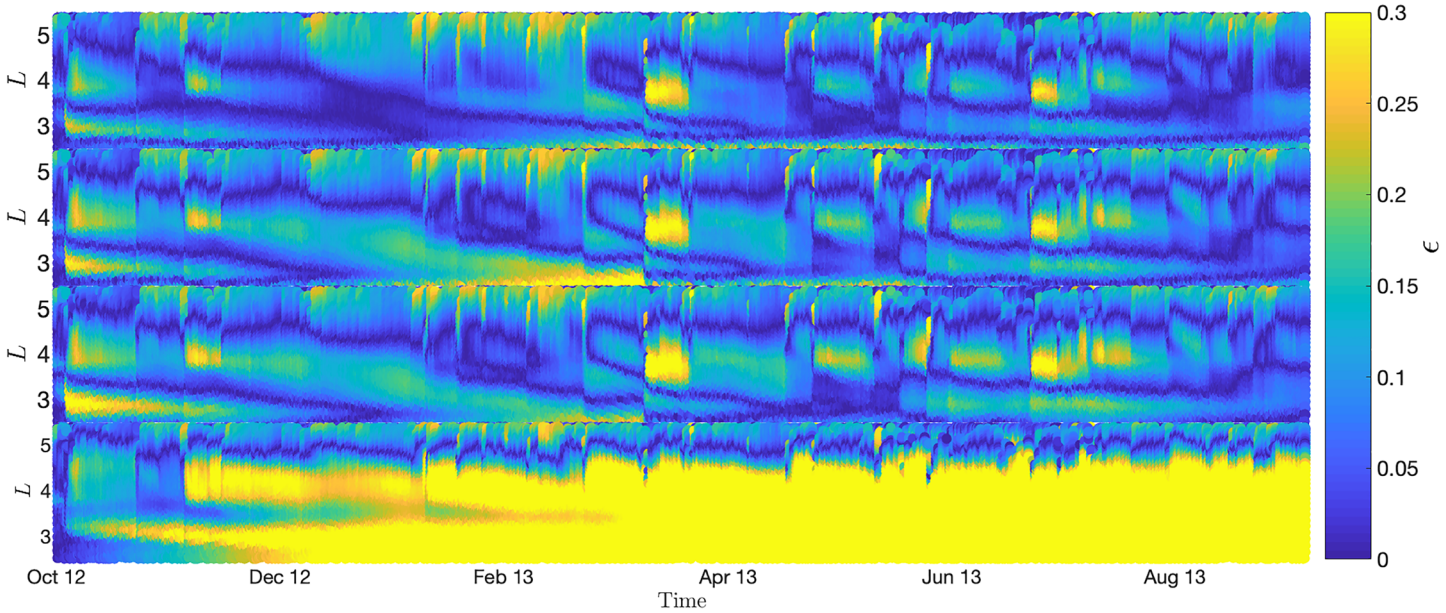

Figure 8. Scatter plot of relative error ϵ in $L - t$ domain given by equation (4), in estimation of phase space density for the time period from October 2012 to September 2013 based on model prediction by equation (1) and parameterized by the maximum a posteriori estimate of posteriors (top panel), reference (Drozdov et al., 2017) (second panel), equation (10) from Ozeke et al. (2014) (third panel), and equation (11) from Ali et al. (2016) (bottom panel).

also exhibits similar error behavior. As expected, the model from Ali et al. (2016) does not perform well at lower values of L. Overall, in terms of the average relative error $\epsilon_{avg} = \sum_i \epsilon_i / m$, where m is the number of Van Allen Probes measurements, it is found that $\epsilon_{avg} = 0.067$, 0.074, 0.078, and 0.326 with the MAP based, reference parameterization, based on Ozeke et al. (2014) and Ali et al. (2016), respectively. It can thus be concluded that at $\mu = 700$ MeV/G and $K = 0.0019$ $G^{0.5}$·Re, the MAP-based model performs better than all the compared models and a small improvement is obtained if the MAP-based parameter values are employed in a deterministic setting.

5. Conclusion

In this paper, we demonstrated a probabilistic framework for estimating the PSD of energetic particles in the magnetosphere, following a standard quasi-linear radial diffusion equation. The probabilistic treatment is based on the Bayesian framework, where, uncertainties are assumed on the parameterization of the diffusion and loss terms, which are iteratively updated using the Van Allen Probes measurements as ground truth. Finally, the informed posteriors are propagated to obtain a probabilistic estimate of particles' phase space density. It is shown that the data is able to reduce the uncertainties considerably, and the probabilistic treatment improves the quality of the predictions. This approach provides an alternate approach to understanding the 1D radial diffusion, since a deterministic form of parameterization may not be able to represent the nonlinear behavior of this boundary-dependent problem.

In terms of the choice of the prior, Ali et al. (2016) obtained uncertainty bounds and attempted to define the probability distribution of the total diffusion coefficient. This could be useful to replace the uniform prior in future investigations, in order to define the domain and improve the convergence of the Bayesian model. The Bayesian estimator predicted longer timescales for the lifetime parameter, to take into account the local acceleration, which is omitted in the 1D model. In the future, a similar approach can be extended to 3D modeling that includes calculation of loss and local acceleration from physics. That would help better quantify the relatively unknown radial diffusion rates. Inferior radial diffusion rate is a very challenging task. In situ measurements do not allow for inferring the m-numbers of waves or observing global distribution of waves in Magnetic Local Time (MLT) and L. Ground measurements allow to find more global maps but observations on the ground may be very inaccurate during disturbed conditions, as ionosphere may shield the waves and they may also be guided to another location, and it may be difficult to trace back the ground observations into space. The Bayesian parameter estimation certainly has a great potential to shed light on this important problem.

Acknowledgments
The *Kp* dataset is obtained from the Omniweb database (https://omniweb.gsfc.nasa.gov). The data generated for this research is available online at https://doi.org/10.4121/uuid:399b9a42-f07e-4b30-a4aa-86abfeeddac0. This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 776262 (AIDA).

References

Albert, J. M., Meredith, N. P., & Horne, R. B. (2009). Three-dimensional diffusion simulation of outer radiation belt electrons during the 9 October 1990 magnetic storm. *Journal of Geophysical Research*, *114*, A09214. https://doi.org/10.1029/2009JA014336

Ali, A. F., Malaspina, D. M., Elkington, S. R., Jaynes, A. N., Chan, A. A., Wygant, J., & Kletzing, C. A. (2016). Electric and magnetic radial diffusion coefficients using the Van Allen probes data. *Journal of Geophysical Research: Space Physics*, *121*, 9586–9607. https://doi.org/10.1002/2016JA023002

Blake, J. B., Carranza, P. A., Claudepierre, S. G., Clemmons, J. H., Crain, W. R., Dotan, Y., et al. (2013). The magnetic electron ion spectrometer (MagEIS) instruments aboard the radiation belt storm probes (RBSP) spacecraft. *Space Science Reviews*, *179*(1), 383–421. https://doi.org/10.1007/s11214-013-9991-8

Bourdarie, S., & Maget, V. (2012). Electron radiation belt data assimilation with an ensemble Kalman filter relying on the Salammbô code. *Annales Geophysicae*, *30*, 929–943. https://doi.org/10.5194/angeo-30-929-2012

Brautigam, D. H., & Albert, J. M. (2000). Radial diffusion analysis of outer radiation belt electrons during the October 9, 1990, magnetic storm. *Journal of Geophysical Research*, *105*(A1), 291–309. https://doi.org/10.1029/1999JA900344

Brooks, S., Gelman, A., Jones, G., & Meng, X.-L. (2011). *Handbook of Markov chain Monte Carlo*. New York: CRC Press.

Camporeale, E. (2019). The challenge of machine learning in space weather: Nowcasting and forecasting. *Space Weather*, *17*, 1166–1207. https://doi.org/10.1029/2018SW002061

Camporeale, E., Chu, X., Agapitov, O., & Bortnik, J. (2019). On the generation of probabilistic forecasts from deterministic models. *Space Weather*, *17*, 455–475. https://doi.org/10.1029/2018SW002026

Camporeale, E., Shprits, Y., Chandorkar, M., Drozdov, A., & Wing, S. (2016). On the propagation of uncertainties in radiation belt simulations. *Space Weather*, *14*, 982–992. https://doi.org/10.1002/2016SW001494

Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, *49*(4), 327–335. https://doi.org/10.1080/00031305.1995.10476177

Drozdov, A. Y., Shprits, Y. Y., Aseev, N. A., Kellerman, A. C., & Reeves, G. D. (2017). Dependence of radiation belt simulations to assumed radial diffusion rates tested for two empirical models of radial transport. *Space Weather*, *15*, 150–162. https://doi.org/10.1002/2016SW001426

Fei, Y., Chan, A. A., Elkington, S. R., & Wiltberger, M. J. (2006). Radial diffusion and MHD particle simulations of relativistic electron transport by ULF waves in the September 1998 storm Journal of Geophysical Research, *111*, A12209. https://doi.org/10.1029/2005JA011211

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC.
Gu, X., Shprits, Y. Y., & Ni, B. (2012). Parameterized lifetime of radiation belt electrons interacting with lower-band and upper-band oblique chorus waves. *Geophysical Research Letters*, *39*, L15102. https://doi.org/10.1029/2012GL052519
Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109. https://doi.org/10.1093/biomet/57.1.97
Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(3), 425–464.
Kennel, C., & Engelmann, F. (1966). Velocity space diffusion from weak plasma turbulence in a magnetic field. *The Physics of Fluids*, *9*(12), 2377–2388.
Kim, K., Shprits, Y., Subbotin, D., & Ni, B. (2011). Understanding the dynamic evolution of the relativistic electron slot region including radial and pitch angle diffusion. *Journal of Geophysical Research*, *116*, A10227. https://doi.org/10.1029/2011JA016684
Kletzing, C. A., Kurth, W. S., Acuna, M., MacDowall, R. J., Torbert, R. B., Averkamp, T., et al. (2013). The electric and magnetic field instrument suite and integrated science (EMFISIS) on RBSP. *Space Science Reviews*, *179*(1), 127–181. https://doi.org/10.1007/s11214-013-9993-6
Kondrashov, D., Shprits, Y., Ghil, M., & Thorne, R. (2007). A Kalman filter technique to estimate relativistic electron lifetimes in the outer radiation belt. *Journal of Geophysical Research*, *112*, A10227. https://doi.org/10.1029/2007JA012583
Lanzerotti, L., Maclennan, C., & Schulz, M. (1970). Radial diffusion of outer-zone electrons: An empirical approach to third-invariant violation. *Journal of Geophysical Research*, *75*(28), 5351–5371.
Li, X. (2004). Variations of 0.7–6.0 MeV electrons at geosynchronous orbit as a function of solar wind. *Space Weather*, *2*, S03006. https://doi.org/10.1029/2003SW000017
Mauk, B. H., Fox, N. J., Kanekal, S. G., Kessel, R. L., Sibeck, D. G., & Ukhorskiy, A. (2013). Science objectives and rationale for the radiation belt storm probes mission. *Space Science Reviews*, *179*(1), 3–27. https://doi.org/10.1007/s11214-012-9908-y
Ozeke, L. G., Mann, I. R., Murphy, K. R., Rae, I. J., & Milling, D. K. (2014). Analytic expressions for ULF wave radiation belt radial diffusion coefficients. *Journal of Geophysical Research: Space Physics*, *119*, 1587–1605. https://doi.org/10.1002/2013JA019204
Ozeke, L. G., Mann, I. R., Murphy, K. R., Rae, I. J., Milling, D. K., Elkington, S. R., et al. (2012). ULF wave derived radiation belt radial diffusion coefficients. *Journal of Geophysical Research: Space Physics*, *117*, A04222. https://doi.org/10.1029/2011JA017463
Reeves, G. D., Chen, Y., Cunningham, G. S., Friedel, R. W. H., Henderson, M. G., Jordanova, V. K., et al. (2012). Dynamic radiation environment assimilation model: Dream. *Space Weather*, *10*, S03006. https://doi.org/10.1029/2011SW000729
Reeves, G. D., McAdams, K. L., Friedel, R. H. W., & O'Brien, T. P. (2003). Acceleration and loss of relativistic electrons during geomagnetic storms. *Geophysical Research Letters*, *30*(10), 1529. https://doi.org/10.1029/2002GL016513
Reeves, G., Spence, H. E., Henderson, M., Morley, S., Friedel, R., Funsten, H., et al. (2013). Electron acceleration in the heart of the Van Allen radiation belts. *Science*, *341*(6149), 991–994.
Roederer, J. G. (1970). *Dynamics of geomagnetically trapped radiation* (Vol. 2). Berlin: Springer-Verlag.
Schulz, M., & Eviatar, A. (1969). Diffusion of equatorial particles in the outer radiation zone. *Journal of Geophysical Research*, *74*(9), 2182–2192.
Schulz, M., & Lanzerotti, L. J. (2012). *Particle diffusion in the radiation belts* (Vol. 7). Berlin: Springer Science & Business Media.
Shprits, Y. Y., Elkington, S. R., Meredith, N. P., & Subbotin, D. A. (2008). Review of modeling of losses and sources of relativistic electrons in the outer radiation belt I: Radial transport. *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(14), 1679–1693. Dynamic Variability of Earth's Radiation Belts https://doi.org/10.1016/j.jastp.2008.06.008
Shprits, Y. Y., Meredith, N. P., & Thorne, R. M. (2007). Parameterization of radiation belt electron loss timescales due to interactions with chorus waves. *Geophysical Research Letters*, *34*, L11110. https://doi.org/10.1029/2006GL029050
Shprits, Y. Y., Subbotin, D. A., Meredith, N. P., & Elkington, S. R. (2008). Review of modeling of losses and sources of relativistic electrons in the outer radiation belt II: Local acceleration and loss. *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(14), 1694–1713. Dynamic Variability of Earth's Radiation Belts https://doi.org/10.1016/j.jastp.2008.06.014
Shprits, Y. Y., & Thorne, R. M. (2004). Time dependent radial diffusion modeling of relativistic electrons with realistic loss rates. *Geophysical Research Letters*, *31*, L08805. https://doi.org/10.1029/2004GL019591
Shprits, Y. Y., Thorne, R. M., Friedel, R., Reeves, G. D., Fennell, J., Baker, D. N., & Kanekal, S. G. (2006). Outward radial diffusion driven by losses at magnetopause. *Journal of Geophysical Research*, *111*, A11214. https://doi.org/10.1029/2006JA011657
Shprits, Y. Y., Thorne, R. M., Reeves, G. D., & Friedel, R. (2005). Radial diffusion modeling with empirical lifetimes: Comparison with CRRES observations. *Annales Geophysicae*, *23*(4), 1467–1471. https://doi.org/10.5194/angeo-23-1467-2005
Spence, H. E., Reeves, G. D., Baker, D. N., Blake, J. B., Bolton, M., Bourdarie,S., et al. (2013). Science goals and overview of the radiation belt storm probes (RBSP) energetic particle, composition, and thermal plasma (ECT) suite on NASA's Van Allen probes mission. *Space Science Reviews*, *179*(1-4), 311–336. https://doi.org/10.1007/s11214-013-0007-5
Stratton, J. M., Harvey, R. J., & Heyler, G. A. (2013). Mission overview for the radiation belt storm probes mission. *Space Science Reviews*, *179*(1), 29–57. https://doi.org/10.1007/s11214-012-9933-x
Su, Z., Xiao, F., Zheng, H., & Wang, S. (2010). STEERB: A three-dimensional code for storm-time evolution of electron radiation belt. *Journal of Geophysical Research*, *115*, A09208. https://doi.org/10.1029/2009JA015210
Su, Z., Zhu, H., Xiao, F., Zong, Q.-G., Zhou, X.-Z., Zheng, H., Wang, Y., Wang, S., Hao, Y.-X., Gao, Z., He, Z., Baker, D. N., Spence, H. E., Reeves, G. D., Blake, J. B., &Wygant, J. R. (2015). Ultra-low-frequency wave-driven diffusion of radiation belt relativistic electrons. *Nature Communications*, *6*, 10096. https://doi.org/10.1038/ncomms10096
Subbotin, D. A., & Shprits, Y. Y. (2009). Three-dimensional modeling of the radiation belts using the Versatile Electron Radiation Belt (VERB) code. *Space Weather*, *7*, S10001. https://doi.org/10.1029/2008SW000452
Summers, D., Ni, B., & Meredith, N. P. (2007). Timescales for radiation belt electron acceleration and loss due to resonant wave-particle interactions: 2. Evaluation for VLF chorus, ELF hiss, and electromagnetic ion cyclotron waves. *Journal of Geophysical Research*, *112*, A04207. https://doi.org/10.1029/2006JA011993
Tu, W., Cunningham, G., Chen, Y., Henderson, M., Camporeale, E., & Reeves, G. (2013). Modeling radiation belt electron dynamics during GEM challenge intervals with the DREAM3D diffusion model. *Journal of Geophysical Research: Space Physics*, *118*, 6197–6211. https://doi.org/10.1002/jgra.50560
Tu, W., Li, X., Chen, Y., Reeves, G. D., & Temerin, M. (2009). Storm-dependent radiation belt electron dynamics. *Journal of Geophysical Researc*, *114*, A02217. https://doi.org/10.1029/2008JA013480
Ukhorskiy, A., & Sitnov, M. (2012). Dynamics of radiation belt particles, *The Van Allen Probes Mission* (pp. 545–578). Boston, MA: Springer.
Walt, M. (1970). Radial diffusion of trapped particles. In B. M. McCormac (Ed.), *Particles and Fields in the Magnetosphere* (pp. 410–415). Dordrecht: Springer Netherlands.

Welling, D., Koller, J., & Camporeale, E. (2012). Verification of SpacePY's radial diffusion radiation belt model. *Geoscientific Model Development*, *5*(2), 277.

Zhelavskaya, I. S., Shprits, Y. Y., & Spasojević, M. (2017). Empirical modeling of the plasmasphere dynamics using neural networks. *Journal of Geophysical Research: Space Physics*, *122*, 11,227–11,244. https://doi.org/10.1002/2017JA024406