Version of March 28, 2007

Identifying the Radiation Belt Source Region by Data Assimilation

J. Koller[1], Y. Chen[1], G. D. Reeves[1], R. H. W. Friedel[1], T. E. Cayton[1], and J. A. Vrugt [2]

Abstract. We describe how assimilation of radiation belt data with a simple radial diffusion code can be used to identify and adjust for unknown physics in the model. We study the drop-out and the following enhancement of relativistic electrons during a moderate storm on October 25, 2002. We introduce a technique that uses an ensemble Kalman Filter and the probability distribution of the forecast ensemble to identify if the model is drifting away from the observations and to find inconsistencies between model forecast and observations. We use the method to pinpoint the time periods and locations where most of the disagreement occurs and how much the Kalman Filter has to adjust the model state to match the observations. Although the model does not contain explicit source or loss terms, the Kalman Filter algorithm can implicitly add very localized sources or losses in order to reduce the discrepancy between model and observations. We use this technique with multi-satellite observations to determine when simple radial diffusion is inconsistent with the observed phase space densities indicating where additional source (acceleration) or loss (precipitation) processes must be active. We find that the outer boundary estimated by the ensemble Kalman filter is consistent with negative phase space density gradients in the outer electron radiation belt. We also identify that specific regions in the radiation belts ($L^* \approx 5-6$ and to a minor extend also $L^* \approx 4$)where simple radial diffusion fails to adequately capture the variability of the observations, suggesting local acceleration/loss mechanisms.

1. Introduction

The highly energetic electron environment in the inner magnetosphere (around geosynchronous orbit and inward) is very dynamic and undergoes constant changes by acceleration, loss, and transport processes. These processes in the radiation belts are important to understand because dynamic variations in this environment can negatively impact the space hardware that our society increasingly depends on.

It has been known since the late 1960's that radial diffusion is a key mechanism influencing radiation belt dynamics, [e.g. *Cornwall*, 1968; *Fälthammar*, 1968; *Schulz and Lanzerotti*, 1974; *Brautigam and Albert*, 2000; *Hilmer et al.*, 2000]. Recently, new observations and increased monitoring evidenced that other processes play an important role as well [*Reeves et al.*, 1998]. For a review see *Friedel et al.* [2002]; *Brautigam and Albert* [2000]; *Green and Kivelson* [2004]. *Reeves et al.* [2003] show that the net effect of geomagnetic activity on radiation belt dynamics is a delicate balance of acceleration, transport, and losses that can lead to either increased or decreased fluxes or to almost no changes at all.

Our new approach is to extend available techniques of data assimilation that are widely used for other

[1]Space Science and Applications, ISR-1, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

[2]Hydrology, Geochemistry, and Geology, EES-6, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

geophysical systems (meteorology, oceanography, ionosphere) to the radiation belts. The term "data assimilation" was coined in the late sixties by the meteorological community to denote a process in which observations distributed in time are merged together with a dynamical numerical model in order to determine as accurately as possible the state of the atmosphere [*Talagrand*, 1997]. The general purpose of data assimilation is to combine all available information essentially consisting of observations and the physical laws which govern the evolution of the system. The latter are available in practice in the form of a numerical model [see *Talagrand*, 1997; *Daley*, 1997, for an introduction].

While diffusion is an important part of the radiation belt description, eventually a self-consistent representation is necessary that includes ring current development and its interaction with radiation belt particles through whistler chorus, hiss, electromagnetic ion cyclotron waves, and other plasma waves and with the changing geomagnetic field. This paper attempts to lay the foundation for the effort to combine all these processes into a Dynamic Radiation Environment Assimilation Model (DREAM) to understand acceleration, transport, and losses in the radiation belts [*Reeves et al.*, 2005]. DREAM is a Laboratory Directed Research and Development project at Los Alamos National Laboratory. It will develop a space radiation model using extensive satellite measurements, new theoretical insights, global physics-based magnetospheric models, and the techniques of data assimilation.

The techniques of data assimilation complement those of traditional first-principle physical models in several ways. We know that a full physical description of the radiation belts require a complete knowledge of the system such that model and data always agree. However, when the model and observations disagree, we have no way to know which aspects of the model produced the disagreement. DREAM develops both approaches in parallel with an eventual convergence.

In the next Section we will present our model-data framework consisting of a radial diffusion code, data from five different satellites, and the ensemble Kalman filter as the overarching umbrella combining data and model predictions. Section 3 describes the results with data assimiation and Section 4 discusses how the Kalman innovation is actually adding a source and loss term to the radial diffusion equation. In Section 5 we estimate the absent source/loss term with the results from data assimilation and Section 6 compares our results with an identical twin experiment. We dicuss our results in Section 7. The Appendix describes the phase space density data and the uncertainties of data, model, and parameters in more detail.

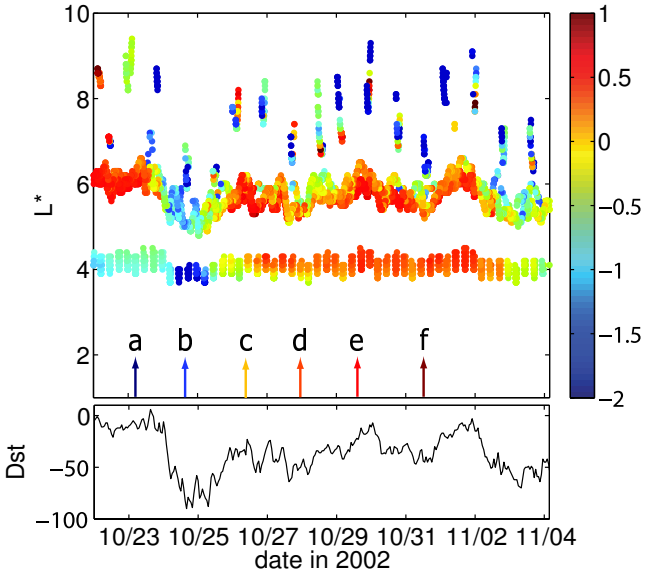

Figure 1. Satellite data and *Dst*. The observations of three LANL-GEO (LANL-97a, 1991-080, 1990-095), Polar, and GPS-ns41 vehicles were converted to phase space densities at constant μ and K. They are plotted here on a normalized log scale. The bottom panel shows *Dst* for the same time period from October 23 to November 4, 2002. This data set is used as input for the ensemble Kalman Filter. We made L^* profile cuts for Figure 4 at the indicated arrows similar to *Green and Kivelson* [2004].

2. The Data and Model Framework

2.1. Data

We used data from three Los Alamos National Laboratory geosynchronous (LANL-GEO) satellites (LANL-97a, 1991-080, 1990-095), Polar, and GPS-ns41 for a specific storm (October 25, 2002) that was chosen by the LWS TR&T (Living with a Star Targeted Research and Technology Program by NASA) Radiation Belt Team (see Figure 1 for phase space density data and disturbance storm-time index *Dst*). This is a moderate storm with a minimum $Dst = -90$ (plotted in the bottom panel of Figure 1). Over the course of eight days, *Dst* slowly recovers but has intermittent minima indicating ongoing activity. Already in Figure 1 we can see radially localized enhancements.

We obtained the phase space densities at given adi-

abatic invariants ($\mu = 2083\,\mathrm{MeV/G}$, $K = 0.1\sqrt{G}\mathrm{R_E}$) by using the angular resolved electron fluxes and local magnetic field magnitude for each of the satellites (see Appendix A). We also applied the global magnetic field configuration from the Tsyganenko 2001 storm-time magnetic field model [*Tsyganenko*, 2002; *Tsyganenko et al.*, 2003]. See *Chen et al.* [2005, 2006] for details on the calculation of phase space densities and adiabatic invariants.

2.2. Radial Diffusion Model

The distribution of relativistic electrons in the radiation belts are described by their phase space density, $f(\mu, J, L^*, t)$ [*Schulz and Lanzerotti*, 1973] where the quantities μ, J, L^* are adiabatic invariants at time t defining the drift motion, periodic gyration and bounce motion of electron in the geomagnetic field [*Roederer*, 1970]. We apply a model that describes only their radial evolution in L^* by using a Fokker-Plank equation with constant adiabatic invariants μ and J

$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{\mathrm{LL}}}{L^2}\frac{\partial f}{\partial L}\right). \qquad (1)$$

We neglect any source or loss terms here. They are simply additive and we will show below that these are included implicitly by the data assimilation algorithm. In fact, identifying where the results after data assimilation deviate from the assumed model is the main focus of this work.

We solve the diffusion equation (1) assuming a discrete meshed grid of dimension N (typically 91 cells) from $1 < L^* < 10$ and use the Crank-Nicolson scheme [*Crank and Nicolson*, 1947] which is an implicit, numerically stable method that does not need to satisfy the Courant condition [*Press et al.*, 1986]. We use a parameterized form of the diffusion coefficient that is a function of magnetic activity [*Brautigam and Albert*, 2000]

$$D_{\mathrm{LL}}(Kp, L) = 10^{(0.506Kp - 9.325)} L^{10}. \qquad (2)$$

The inner boundary at $L^* = 1$ is fixed at zero, however, the outer boundary is a free parameter that can be adjusted by our data assimilation algorithm.

The initial condition for the grid is a steady state vector that has been calculated with a constant Kp over a very long time and outer boundary $f_b = 1$. The steady state is then simply multiplied by a factor to match the very first data point. Also, all data has been scaled by a global factor to obtain an average $\langle f \rangle = 1$.

2.3. Ensemble Kalman Filter

The term "data assimilation" is short for "model-based assimilation of observations", i.e. data assimilation is the combination of a given physical model with observations. The purpose is to find the most likely estimation to the true state (which is unknown) using the information provided by the chosen physical model and the available observational data considering both of their uncertainties and the limitations of both model and observations. Data assimilation methods are based on, and can be derived from, Bayesian statistics, minimum variance, maximum likelihood, or least square methods [*Maybeck*, 1979; *Kalnay*, 2003; *Daley*, 1991; *Talagrand*, 1997; *Tarantola*, 1987; *Tarantola and Valette*, 1982].

One popular method for data assimilation is the Kalman Filter [*Kalman*, 1960]. It is an optimal recursive data processing algorithm [*Maybeck*, 1979, p. 4] that has become a favorite for many engineering application including the navigational system on the Apollo mission, GPS stand-alone devices, and many more [*Sorenson*, 1985].

We combine the phase space densities at each grid point into a single vector $\mathbf{x}$, called the state vector, and the observations at time t_i into the observational state or data vector $\mathbf{y}^{\mathrm{o}}(t_i)$. The Kalman Filter method can be summarized in three steps which are illustrated in Figure 2.

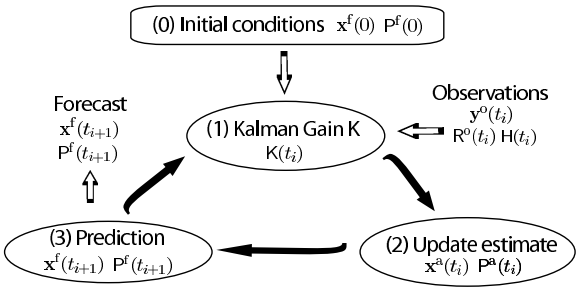

Figure 2. Flow diagram of the recursive Kalman filter algorithm. Starting with initial state vector $\mathbf{x}^{\mathrm{f}}(0)$ and an estimation of its error in $\mathsf{P}^{\mathrm{f}}(0)$, the first step is to compute the Kalman gain matrix K. Then observations $\mathbf{y}^{\mathrm{o}}(t_i)$ are used to calculate the state estimate $\mathbf{x}^{\mathrm{a}}(t_i)$. Step three yields the forecast state vector $\mathbf{x}^{\mathrm{f}}(t_{i+1})$ which is used as input for the next cycle.

(1) Gain computation: which yields the "Kalman

gain matrix" or "weight matrix"

$$\mathsf{K}(t_i) = \mathsf{P}^{\mathrm{f}}(t_i)\mathsf{H}^{\mathrm{T}}(t_i)\left[\mathsf{H}(t_i)\mathsf{P}^{\mathrm{f}}(t_i)\mathsf{H}^{\mathrm{T}}(t_i) + \mathsf{R}^{\mathrm{o}}(t_i)\right]^{-1} \quad (3)$$

which depends on the error covariance of the current forecast P^{f} and the observational uncertainty R^{o}. The operator H projects the model space into the observational state space, i.e. it pulls out the specific grid points from the state vector where observations are made. H^{T} indicates the transpose of H.

(2) State estimate: which uses the Kalman gain K to weight the "observational residual" (in the older meteorological literature) or the "innovation vector" $\mathbf{d} = \mathbf{y}^{\mathrm{o}} - \mathsf{H}\mathbf{x}^{\mathrm{f}}$ and computes the "state estimate" or "assimilated state" $\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{f}} + \mathsf{K} \cdot \mathbf{d}$.

(3) State forecast or prediction: The next step is to apply a "forward model operator" M which results in the "forecast state vector" $\mathbf{x}^{\mathrm{f}}(t_{i+1}) = \mathsf{M}[\mathbf{x}^{\mathrm{a}}(t_i)]$ that can be compared with new observations at time t_{i+1} in the next cycle.

The steps above also include the calculation of the error covariances P^{f} and P^{a}. In the standard linear Kalman filter they are explicitly calculated [*Koller et al.*, 2005]. However, we decided to use an ensemble Kalman Filter [*Evensen*, 1994, 2003] instead, which is a variant of the classical Kalman filter as described above. The standard Kalman filter can only handle linear problems but we have extended our state vector by model parameters which make the model non-linear.

One can describe the error statistics from non-linear models by using a Monte-Carlo technique like the ensemble Kalman filter. The ensemble members are created by randomly perturbing the state vector, separately advancing them in time by using the model, and then comparing them to each other. The new, most likely, forecast is the mean of the whole forecast ensemble. The spread of the forecast ensemble members Δ_e determines the uncertainty of the forecast. The ensemble covariance matrixes around the ensemble mean $\overline{\mathbf{x}}$ are defined as

$$\mathsf{P}^{\mathrm{f}} = \overline{(\mathbf{x}^{\mathrm{f}} - \overline{\mathbf{x}^{\mathrm{f}}})(\mathbf{x}^{\mathrm{f}} - \overline{\mathbf{x}^{\mathrm{f}}})^{\mathrm{T}}} \quad (4)$$

$$\mathsf{P}^{\mathrm{a}} = \overline{(\mathbf{x}^{\mathrm{a}} - \overline{\mathbf{x}^{\mathrm{a}}})(\mathbf{x}^{\mathrm{a}} - \overline{\mathbf{x}^{\mathrm{a}}})^{\mathrm{T}}} \quad (5)$$

where now the overlines denote an average over the ensemble. Thus, we can use an interpretation where the ensemble mean is the best estimate and the spreading of the ensemble around the mean is a natural definition of the error in the ensemble mean. Increasing the number of ensemble members N_e gives a better resolution of the probability distribution of the state vector and the error of the sampling decreases proportional to $1/\sqrt{N_e}$ [*Evensen*, 2003]. We test the convergence of our results with different numbers of ensembles and find that for $N_e > 30$ all our results and conclusions are consistent.

We use a so-called "augmented state vector" approach [*Lainiotis*, 1971; *Ljung*, 1979] where the state vector is extended by parameters of the physics model. We added the phase space density of the outer boundary f_b to the state vector which then reads

$$\mathbf{x}^e = [\mathbf{x},\ f_b]^T. \quad (6)$$

The diffusion equation (1) by itself is linear but since we have augmented the state vector with parameters that are used in the solution it becomes a non-linear equation $\mathbf{x}^{i+1} = \mathsf{M}(\mathbf{x}^i)$. All of the physics assumed in our model is contained in the matrix M which simply advances the state from time t_i to t_{i+1}. The matrix M will be used for the prediction part in the Kalman Filter.

3. Data Assimilation Results

Assimilating and combining all data with our 1-D radial diffusion code results in the states shown in Figure 3. The whole data set and, hence, also the resulting phase space density is normalized by a fixed value. The first two days (marked gray in Figure 3) before the strong drop-out in phase space density should be considered as an adjustment period for the Kalman Filter because of the initial conditions which were simply a steady state system. Also, the outer boundary was still at a high value from the initial steady state system. The Kalman Filter assimilated two days worth of data to adjust from the initial conditions to the observed radiation belt profile. The boundary consistently dropped during that same time period and stayed low for the rest of the studied interval. One can also see in Figure 3 that the outer boundary is modulated mostly by Polar data due to the proximity of measurement locations to the computational boundary at $L^* = 10$.

After the minimum Dst on October 25, the phase space density dramatically rises within just two days between $L^* = 4 - 7$. After a maximum is reached, Figure 3 shows several smaller decreases. We find that the evolution of the phase space density is well correlated with the temporal evolution of Dst.

Compared to Figure 1, the entire L^*-time space is now filled (top panel in Figure 3). However, it is important to keep in mind that the uncertainty of each

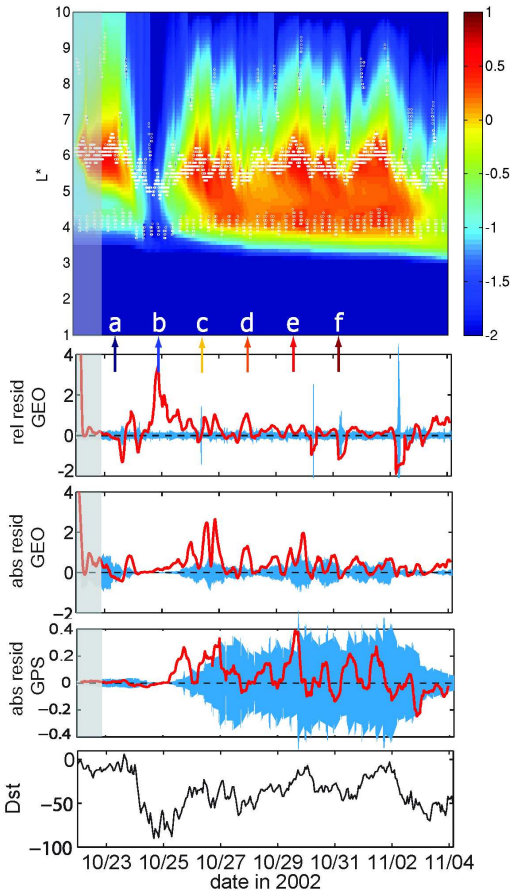

Figure 3. (1) Assimilated state result using data from Figure 1 and a 1-D diffusion with variable outer boundary. The locations of satellite data that went into this result are shown as white circles. Cuts for density profiles in Figure 4 are indicated as well. (2) Relative residual used by the ensemble Kalman Filter to adjust for model discrepancies. The light blue area depicts the relative uncertainty of the forecasts. The red line shows the relative residual that is positive for a long period which results in an increase of the phase space densities to match the observations of LANL-GEO satellite 1991-080. We used a moving average of 3.3 hours. (3) Same as panel 2 but showing the absolute residual. (4) Absolute residual at GPS orbit. Note the different scale compared to panel two and three. The gray area in panel one to four indicates the initial adjustment period for the data assimilation algorithm. (5) *Dst* for the same time period.

forecast increases with increasing distance to the data: The further one goes away from regions where data is available, the lower the confidence in the prediction of

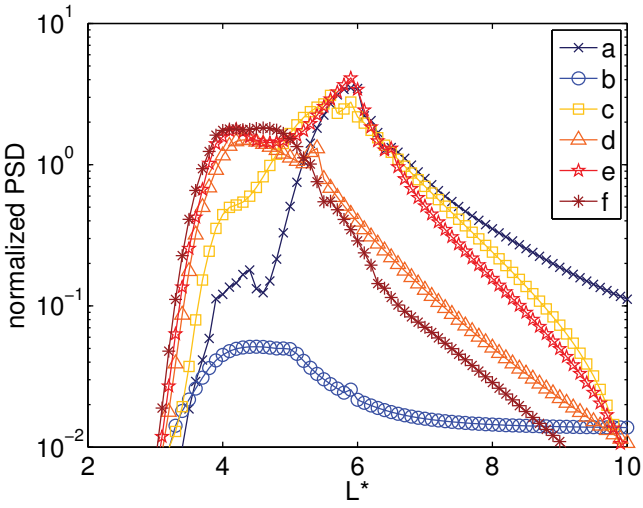

Figure 4. Phase space density profiles of the assimilated states in Figure 3. Cut times are indicated by the letter a-f.

the state will be.

Figure 4 shows cuts through the phase space density at six different times which are indicated in Figure 3. All cuts from pre-storm through the recovery phase show a negative slope for $L^* > 6$ and either a single peak or a double peak between $4 < L^* < 6$ indicating a source process in the inner regions of the radiation belts and not radial inward diffusion which is consistent with the results of *Green and Kivelson* [2004].

Looking at Figure 3 one might wonder how a model without a source or loss process can produce such distinctive phase space density peak in the center of the radiation belts. The reason lies in the data assimilation as described below.

4. Kalman Innovation Adding Source/Loss Processes

The second step in the Kalman filter where the state estimate is calculated warrants a more detailed discussion because this is where a source/loss term is effectively incorporated into the Kalman Filter. The solution to the diffusion equation (1) can be formulated as $\mathbf{x}^{\mathrm{f}}(t_{i+1}) = \mathsf{M}\mathbf{x}^{\mathrm{a}}(t_i)$. Replacing $\mathbf{x}^{\mathrm{f}}$ in the state analysis equation $\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{f}} + \mathsf{K} \cdot \mathbf{d}$ yields

$$\mathbf{x}^{\mathrm{a}}(t_{i+1}) = \mathsf{M}\mathbf{x}^{\mathrm{a}}(t_i) + \mathbf{S} \tag{7}$$

where $\mathbf{S} = \mathsf{K} \cdot \mathbf{d}$ is acting as the source term. This is equivalent to solving

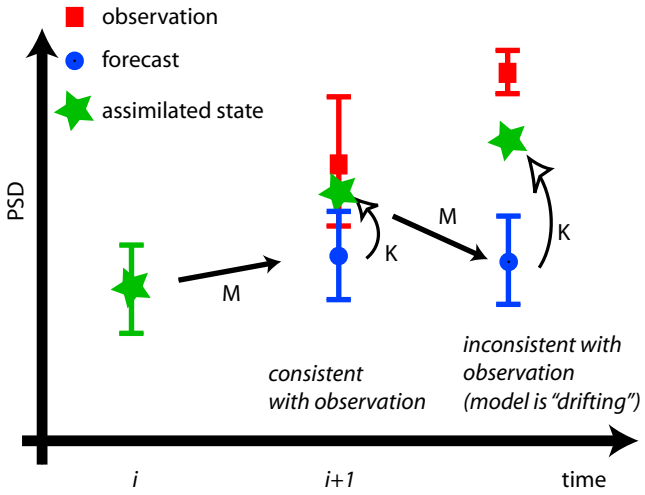

Figure 5. Ensemble Kalman filter diagram explaining how the Kalman innovation is implicitly acting as a source or loss process. The current state forecast (blue circle) is compared to the observation (red box). The Kalman gain is calcualted as a function of the uncertainties. If the forecast is within the errors of the observations, the forecast model adequately describes the data. However, if the observation falls outside of the forecast uncertainty, the model has a large discrepancy and the Kalman filter will apply a significant amount of source or loss in the form of $\mathbf{S} = \mathsf{K} \cdot \mathbf{d}$ in order to match the observation. The result is the assimilated state. We use large values of $\mathbf{S}$ to indicate strong discrepancies between the model forecast and the observations.

$$\frac{\partial f}{\partial t} = L^2 \frac{\partial}{\partial L}\left(\frac{D_{\mathrm{LL}}}{L^2}\frac{\partial f}{\partial L}\right) + \widetilde{S} \qquad (8)$$

but with $\widetilde{S}$ as a function of time and the radial coordinate L^*. However, $\mathsf{K} \cdot \mathbf{d}$ is a full adjustment to the model state and applied in the Kalman filter during a comparison of forecast with an observation whereas $\widetilde{S}$ is applied with each time step.

The magnitude of the source or loss $\mathbf{S}$ in Equation (7) depends on the uncertainty of model and observations and how they compare to each other. If the confidence in the observations is low, the estimate will favor the model with small values in K. The elements in the source vector $\mathbf{S} = \mathsf{K} \cdot \mathbf{d}$ will then be close to zero. On the other hand, if the uncertainty of the model is large, then more weight will be given to the observations with large values in K. If K is large and the difference between the observation and the forecast $\mathbf{y}^{\mathrm{o}} - \mathsf{H}\mathbf{x}^{\mathrm{f}}$ is also large, then elements in $\mathbf{S}$ will be large as well. See Appendix B for a discussion on model and data uncertainties.

Since elements in $\mathbf{S}$ can be positive or negative, the exact same arguments apply for losses. We note, however, that $\widetilde{S}$ in Equation (8) represents overall net sources or losses at a given time step.

We introduce here a method that can be used to identify time intervals where disagreements between the model and the observations imply that a simple diffusion model (without additional source/acceleration or loss/precipitation terms) is inadequate to describe the dynamics of the system. Every time an ensemble member is integrated from time t_i to t_{i+1} it becomes a "forecast" (this term is used regardless of whether t_{i+1} is in the future or applied to retrospective data sets). We employ the forecast ensemble to calculate the average forecast and the spread of the ensemble describing the probability distribution of that average forecast (Figure 5). The mean forecast is compared to the observations using the innovation equation $\mathbf{d} = \mathbf{y}^{\mathrm{o}} - \mathsf{H}\mathbf{x}^{\mathrm{f}}$. If the mean forecast plus or minus its uncertainty (based on the ensemble) overlap the uncertainty of the observation, then the forecast can be considered as consistent with the observations. But if the forecast by the model is outside of the observational error bar, the model was not able to predict the observation well enough. See Figure 5 for a sketch of this process.

If the model is frequently too low or too high, the model is said to be "drifting" relative to the observations. When this occurs the Kalman Filter responds by adjusting the current state $\mathbf{x}^{\mathrm{a}}$ away from the model forecast $\mathbf{x}^{\mathrm{f}}$ and toward the observations $\mathbf{y}^{\mathrm{o}}$. In this case it does that by adding or subtracting phase space densities - essentially mimicking the effects of source or loss terms by the means of the "Kalman innovation" term $\mathsf{K} \cdot \mathbf{d}$. We can use this term to estimate the missing source term in Equation (1) and localize the L^*-shell region. We will discuss this in the next section.

5. Estimating the Missing Source/Loss

The second panel in Figure 3 shows the relative residual d_j/x_j^f (red line) between the forecast and the observations of one Los Alamos National Laboratory geosynchronous (LANL-GEO) satellite 1991-080. x_j^f is short for the predicted phase space density at the corresponding L^* grid point of satellite j or in a more mathematical way $x_j^f := \mathsf{H}_j\mathbf{x}^{\mathrm{f}}$. The operator H_j is the j-th row of the observation operator H. The relative residual is positive (negative) when the observations are significantly higher (lower) than the forecast and hence the Kalman Filter is adding (subtracting) phase space den-

sity to compensate and to match the observations. The relative residual is plotted in Figure 3 with a moving average of 3.3 hours. Note that the peak in the relative residual d_j/x_j^f of satellite j does not correspond to the peak in phase space density. This is because the phase space density is low. It is the relative variation that is large. For comparison, the absolute residual d_j (third panel in Figure 3 correlates better with the peaks in phase space density. The light blue area indicates a one sigma uncertainty of the prediction $P_j^f(t)^{0.5}$ at the L^* location of satellite j. If the red line overlaps with the blue area, the residuals are within the prediction uncertainty and noise of the observations.

This type of plot can be used to identify when and where "missing physics" in the model becomes apparent. In our case, the missing source/loss term in the radial diffusion model is most obvious between October 25-27 for geosynchronous satellites (Figure 3), where the model is significantly drifting away from the observations. That trend is compensated by $\mathsf{K}\cdot\mathbf{d}$ in the state analysis equation $\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{f}} + \mathsf{K}\cdot\mathbf{d}$ which can also be interpreted as a source/loss term (see Section 4 for a more complete discussion).

We compare the residuals at LANL-GEO to the residuals at GPS and find that they are a factor of 10 larger for geosynchronous satellites (Figure 3 panel 4). The Kalman Filter added only a much smaller amount to the phase space density. This indicates that the real source region of acceleration is not at the GPS orbit but rather at geosynchronous orbit or between GPS and LANL-GEO.

The dimension of the Kalman innovation $\mathsf{K}\cdot\mathbf{d}$ is the same as the state vector and can be used to identify where the Kalman Filter has added the most phase space density. We sum over all Kalman innovations $\mathsf{K}\cdot\mathbf{d}$ during the recovery phase from October 25 to November 2, 2002 and find the largest amount is added between $L^* = 5-6$ (Figure 6). This indicates that the phase space density measured by GPS around $L^* = 4$ can mostly be accommodated by inward radial diffusion. There is a small peak of added source but it is a factor 10 smaller than at geosynchronous orbit. Regions outside of $L^* = 6$ are also not receiving any additional source. Changes in phase space density outside of $L^* \approx 6$ are consistent with outward radial diffusion.

This particular finding is not necessarily the same for other storms. We studied here the radiation belt response due to a fairly standard interaction region with a sector reversal in the solar wind. Other, more extreme storms, could very well lead to a different response and a different localized particle acceleration region as *Shprits et al.* [2006] suggest an $L^* \approx 3$ during the Halloween storm in 2003.

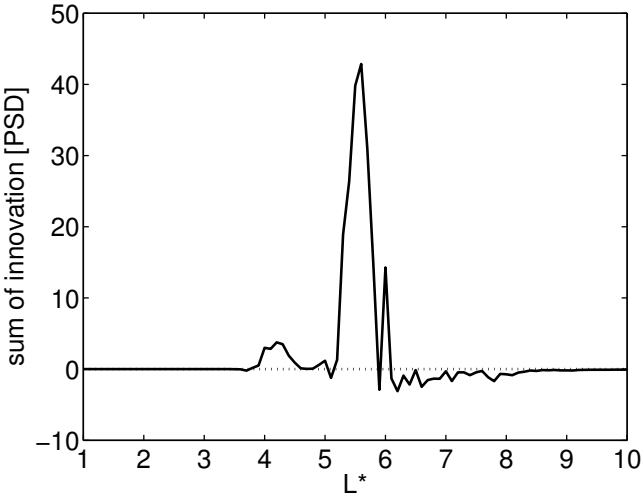

Figure 6. Kalman innovations $\mathsf{K}(\mathbf{y}^{\mathrm{o}} - \mathsf{H}\mathbf{x}^{\mathrm{f}})$ versus L^* summed over the time period of the recovery phase from October 25 to November 2, 2002. This shows at what L^*-shells the ensemble Kalman Filter added phase space densities like a source term in order to match the observations. Most of the source is added between $L^* = 5-6$ whereas the enhanced phase space density outside this region is mostly explained by radial diffusion.

We also note that the ensemble Kalman Filter did not raise the boundary condition to facilitate inward diffusion but rather added phase space densities locally like a source term otherwise the forecasts would become inconsistent with Polar.

The next step, but beyond this paper, will be to identify the physical processes that are responsible for the model trend and to include them in a new version. That should greatly reduce the residuals and lead to a better model based on better physical understanding. Possible candidates for these model trends are most likely wave-particle interactions (whistler chorus, hiss, electromagnetic ion cyclotron waves, etc.) which are also being studied by other modeling and data analysis efforts [*Horne and Thorne*, 1998; *Meredith et al.*, 2002; *Summers*, 2005; *Varotsou et al.*, 2005; *Shprits et al.*, 2005; *Green and Kivelson*, 2004].

6. Discussion and Conclusion

The advantage of an analysis with data assimilation lies in obtaining a complete picture of the radiation belt, estimating free parameters like the outer boundary, and

incorporating uncertainty in data and model.

We studied the combination of a 1-D radial diffusion code with an ensemble Kalman Filter and assimilated data from 5 satellites for the time period from October 22 to November 4, 2002. The data from three LANL-GEO, Polar, and GPS show strong enhancements after a drop out. We used the ensemble Kalman Filter with a state vector that was extended to allow the outer boundary to adjust freely. We find that the outer boundary stayed low during the whole time period indicating that a local acceleration process is dominating the dynamics instead of an inward radial diffusion from the boundary.

We were also using the ensemble Kalman Filter to identify time periods when the model is drifting away from the observations suggesting diffusion alone without internal sources and losses provides an incomplete physical description of the dynamics. The specific equations of the Kalman Filter can compensate for such "missing physics". We find that the largest relative residual between the forecast ensemble and the observations are at the minimum of Dst. This entails that the actual phase space density is a factor 3-4 larger than the predicted value. That relative residual is falling off as the phase space density rises but stays positive until October 26. The absolute residual corresponds to the peaks in phase space density of the assimilated state around October 27 and 30.

We find that the source region may be very localized. The Kalman Filter adds most of the phase space density between $L^* = 5 - 6$. Radial diffusion then redistributes the effects of source or loss. It is possible that this results is due to the location of geosynchronous observations. The new RBSP (Radiation Belt Storm Probe) mission with its geo-transfer orbit will help to pin down the answer.

We also want to point out that we used the diffusion coefficients from *Brautigam and Albert* [2000] assuming they were fixed. However, *Brautigam and Albert* [2000] point out in their paper that these diffusion coefficients have a large uncertainty. Also, more recent studies by *Fei et al.* [2006] find that $D_{\mathrm{LL}} = D_0 L^n$ where $D_0 = 1.5 \times 10^{-6} \mathrm{days}^{-1}$ and $n = 8.5$ representing an average of the time-varying ULF-driven diffusion coefficient. We could have added the diffusion coefficients to the state vector as well but we will leave this to a future study.

We plan to repeat the same analysis when we have models with explicit sources and losses to quantify the effect on residuals. In a next step we will identify the processes behind these explicit sources and losses and include them in a new model which should have much smaller residuals.

In summary, the ensemble Kalman Filter can be used with a relatively simple model to identify where and when sources and losses operate. This makes data assimilation a promising method to study radiation belt data to get a better understanding of the tug-of-war between physical processes causing acceleration, losses, and transport.

Appendix A: Phase Space Density Data

On board of the Los Alamos National Laboratory geosynchronous (LANL-GEO) satellites (1990-095, 1991-080, and LANL-97A), the Synchronous Orbit Particle Analyzer (SOPA) instrument [*Belian et al.*, 1992] can measure the full three-dimensional electron distribution from 50 keV to more than 1.5 MeV in each spin. Since the LANL-GEO satellites carry no magnetometer instruments, we employ the method developed by *Thomsen et al.* [1996] through which the local magnetic field direction can be derived from the measurement of the plasma distribution by another instrument on board - the Magnetospheric Plasma Analyzer (MPA), to obtain the pitch angle distribution [*Chen et al.*, 2005]. In this work the LANL-GEO electron data have a 10 minute time resolution, and we use the empirical magnetic field model to calculate the adiabatic invariants (μ, J, L^*).

The Polar satellite, with a polar orbit of $2 \times 9 R_E$, crosses the magnetic equatorial plane every 18 hours just outside of GEO during the time periods studied here. The Comprehensive Energetic Particle and Pitch Angle Distribution (CEPPAD) experiment [*Blake et al.*, 1995] on board of Polar provides angular resolved flux data of energetic electrons, covering the energy range from 30 keV-10 MeV. Flux data have a time resolution of 3.2 min, and we only use the measurements at the apogee equatorial crossings. Polar also carries a Magnetic Field Experiment (MFE) [*Russell et al.*, 1995] measuring magnetic field vectors. Therefore, μ can be calculated directly but J and L^* still require the model.

The GPS satellites have a circular orbit with a radius of $4 R_E$ and inclination of 55 degrees, which makes them cross the equatorial plane every 6 hours. The electron data used in this work are from one satellite, GPS-ns41, measured with the BBD-IIR (Burst Detector Dosimeter IIR) obtaining differential energy electron fluxes from 77 keV up to > 5 MeV [*Cayton et al.*, 1998]. The flux data have a time resolution of 4 min. Since GPS satellites are three-axis stabilized and have no magnetometer on board, we assume here an isotropic pitch angle distribution and use the Tsyganenko 2001 storm-time

magnetic field model [*Tsyganenko*, 2002; *Tsyganenko et al.*, 2003] for obtaining (μ, J, L^*).

We made strong efforts to calibrate the measured data between satellites. The inter-calibration between the three LANL-GEO satellites was obtained in *Chen et al.* [2005] by matching the phase space densities, that is, comparing the phase space density values of electrons with the same combination of (μ, K, L^*) but measured by satellites at different spatial locations during magnetically quiet times. The same method is applied to obtain the inter-calibration between LANL-GEO and Polar [*Chen et al.*, 2006]. Also, a preliminary inter-calibration between Polar, LANL-GEO, and GPS fluxes was done following the procedure decribed in *Friedel et al.* [2005].

One distinguished feature in Figure 1 is that the L^* positions of satellites vary greatly with time, even during quiet times. This variation involves two parts: (1) The diurnal variation for LANL-GEO satellites, which have nearly fixed equatorial radial distances, is caused by the asymmetric magnetic field. For larger L^* and on the night side, the measured field is more stretched and weaker than on the day side [*Chen et al.*, 2005, 2006]. This variation dominates during quiet time. (2) After the diurnal change is removed, the remaining variation in L^* is more pronounced during storm times and is caused by changing magnetospheric current systems (especially the ring current). These current systems simultaneously cause the change in Dst and therefore lead to the "Dst effect". Electrons move to different spatial position so they conserve the third adiabatic invariant [*Kim and Chan*, 1997]. To conserve the invariants, the "Dst effect" requires the drift shell to move radially outward and consequently leaves the GEO satellite to find itself on a new drift shell with smaller L^* value. The same reason makes the GEO satellites move back to the pre-storm L^* shells in the recovery phase. This mechanism applies to all satellites. Such changes in L^* justify the importance of comparing phase space densities in a correct magnetic coordinate system.

Appendix B: Data, Model, and Parameter Uncertainties

Uncertainties of the observations Δ_y and the model Δ_M are an important ingredients to every data assimilation process except the method of direct insertion where the data is assumed to have no error at all. Also, the task is left to the model to propagate the information from data to other surrounding locations. This is in strong contrast to other data assimilation methods where correlations between all locations are used to find the best approximation to the true state while staying consistent with all data points.

We estimate the observational uncertainty by comparing different satellite measurements of the same parameter against one another and adding an estimate of systematic uncertainties [*Friedel et al.*, 2005]. We use a 1-D grid in L^* and were therefore able to find many conjunctions between geosynchronous satellites. A statistical analysis of the conjunctions gave us the relative uncertainty of the observations. We find a relative uncertainty of 30% using 6500 conjunctions over the course of half a year. We applied then the same uncertainty to all instruments including GPS and Polar that did not have any conjunctions (along a drift shell) with data from LANL-GEO satellites. We note that in practice this is often only a best estimate of the observational uncertainty.

Model uncertainties, Δ_M, are determined by a combination of the ensemble spread, Δ_e, in the ensemble Kalman Filter and free parameters like the outer boundary. They are much more difficult to estimate, especially since we know that our simple 1-D diffusion model is incomplete but we do not know the magnitude of the resulting model uncertainty. This is an still ongoing research topic even in the atmospheric data assimilation community [see *Mitchell et al.*, 2002, and references therein].

We did several tests and find that the ensemble spread should be $\Delta_e \gtrsim \Delta_y$ in order to leave enough room for the Kalman Filter to adjust for the fast changes in phase space densities of the observations. This way, we put enough confidence in the observation so that unknown physical processes in the model are compensated for by the source term which largely depends on the uncertainty of observations versus model. We find that as long as the assumed model uncertainty is approximately equal or larger than the data uncertainty, the results are stable and, moreover, consistent with the data. In any case, the model uncertainty should adequately represent the "large" portion of missing source/loss in the model although the exact number might be difficult to determine.

The uncertainty of the outer boundary is estimated by how fast the observations from Polar satellite (at $L^* \approx 7 - 10$) change within a certain time period.

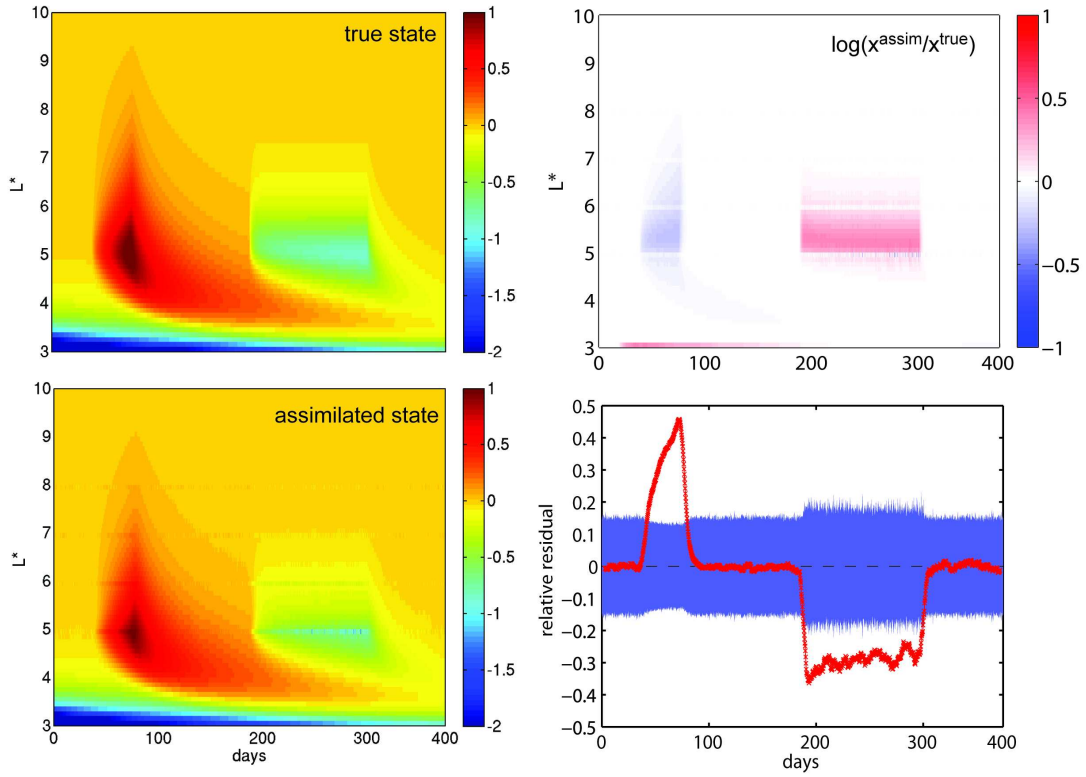

Figure C1. Identical twin experiment with an artificial source region followed by a loss period. The initial conditions are derived from a steady state system with constant boundaries. The true state $\mathbf{x}^{\text{true}}$ is shown in the upper left panel. We created artificial data using the true state and assimilated the data with a model that does not contain source/loss processes. The result of the assimilated state $\mathbf{x}^{\text{a}}$ is shown in the lower left panel. Although the physics model did not contain a source/loss process the Kalman Filter can compensate for it. The upper right panel displays the ratio $\mathbf{x}^{\text{a}}/\mathbf{x}^{\text{true}}$ on a logarithmic scale. The lower right panel shows the relative residual to point out the time periods where the Kalman Filter had to add or subtract a significant amount of phase space density in order to match the observations. The blue area describes the uncertainty of the forecast.

Appendix C: Tests with Identical Twin Experiments

We use "identical twin experiments", a term introduced by the data assimilation community, to test the results with artificial data. We use these experiments to assess their applicability to identify source and loss processes. The method of the identical twin experiment is to create artificial data with simulated uncertainties and to test assimilation schemes in such a controlled environment [*Koller et al.*, 2005; *Naehr and Toffoletto*, 2005]. This has several advantages over using real data. First, we have exact knowledge of the "true" state at all times and all locations. This is something that can never be found in reality because data as well as models, no matter how good they are, are only approximations to the true state. Second, data can be artificially created with any kind of error statistics along arbitrary satellite orbits.

Observations and true state are related by $\mathbf{y}^{\text{o}} = \mathbf{x}^{\text{true}} + \epsilon^{\text{o}}$ where ϵ^{o} is the unknown observational uncertainty but with known statistical properties. Also we will never know the exact true state $\mathbf{x}^{\text{true}}$. However, in the "identical twin experiment" we get to assume a certain "true state" and create an artificial data point $\mathbf{y}^{\text{o}}$ by randomly adding an error reflecting the chosen statistical uncertainty.

For the following identical twin experiment, we create a steady state system with constant boundaries and a constant rate of radial diffusion. We then add an artificial source region followed by a loss period (see upper

left in Figure C1). We sample the true state every six hours with five artificial satellites at constant L^*-shells between $3 < L^* < 8$ and randomly perturb these observations to simulate observational uncertainty. These observations are then fed into the Kalman Filter algorithm for assimilation with a physics model that contains only radial diffusion but no source or loss processes.

The resulting Kalman Filter output, the assimilated state, is shown in Figure C1 (lower left panel). It contains information from the data and the model prediction. We find that even when the model does not contain source/loss processes, the Kalman Filter can compensate model discrepancies by adding (subtracting) phase space density very efficiently in order to match the observations. We identify such model trends or "drifts" by comparing the ensemble of predictions with observations. If the mean of the ensemble forecast members is inconsistent with the observation and their uncertainty, then the Kalman Filter is adding a significant amount by the means of $\mathsf{K} \cdot \mathbf{d}$ in the state analysis equation $\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{f}} + \mathsf{K} \cdot \mathbf{d}$ which can also be interpreted as a source/loss term.

We visualize model discrepancies or the observational residual by plotting components of the innovation vector $\mathbf{d}$. Each component represents the difference between the forecast and the observation of a particular satellite. Specifically, in the lower right panel of Figure C1, we plot a smoothed relative observational residual d_j/x_j^f to eliminate noise. We also plot the uncertainty of the forecast x_j^f at the location of satellite j (blue area). If the residual d_j (red line) of satellite j falls outside of the uncertainty then the model discrepancy is significant. Further, if the residual is positive (negative) then phase space density was added (subtracted) by the Kalman Filter compensating for an unspecified source process (loss process).

This identical twin experiment shows that the Kalman filter can be used to compensate for and identify regions and time periods with significant "unknown" acceleration and loss processes.

References

Belian, R. D., G. R. Gisler, T. Cayton, and R. Christensen (1992), High-Z energetic particles at geosynchronous orbit during the great solar proton event series of October 1989, *J. Geophys. Res.*, *97*, 16,897–16,906.

Blake, J. B., et al. (1995), CEPPAD: Comprehensive energetic particle and pitch angle distribution experiment on Polar, *Space Sci. Rev.*, *71*, 531–562.

Brautigam, D. H., and J. M. Albert (2000), Radial diffusion analysis of outer radiation belt electrons during the October 9, 1990, magnetic storm, *J. Geophys. Res.*, *105*, 291–310, doi:10.1029/1999JA900344.

Cayton, T. E., D. M. Drake, K. M. Spencer, M. Herrin, T. J. Wehner, and R. C. Reedy (1998), Description of the BDD–IIR: Electron and proton sensors on the GPS, *Tech. Rep. LA-UR-98-1162*, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

Chen, Y., R. H. W. Friedel, G. D. Reeves, T. G. Onsager, and M. F. Thomsen (2005), Multisatellite determination of the relativistic electron phase space density at geosynchronous orbit: methodology and results during geomagnetically quiet times., *Journal of Geophysical Research*, *110*(A10), A10,210 – A10,225.

Chen, Y., R. H. W. Friedel, and G. D. Reeves (2006), The phase space density distributions of energetic electrons in the outer radiation belt during two gem im/s selected storms., *Journal of Geophysical Research*, *0*(0), 0.

Cornwall, J. M. (1968), Diffusion processes influenced by conjugate-point wave phenomena, *Radio science*, *3*, 740 –.

Crank, J., and P. Nicolson (1947), A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type, *Proceedings of the Cambridge Philosophical Society*, *43*, 50–67.

Daley, R. (1991), *Atmospheric Data Analysis*, Cambridge University Press, Cambridge.

Daley, R. (1997), Atmospheric data assimilation, *Journal of the Meteorological Society of Japan*, *75*(1B), 319 – 329.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics., *Journal of geophysical research*, *99*(C5), 10,143 – 10,162.

Evensen, G. (2003), The ensemble kalman filter: theoretical formulation and practical implementation., *Ocean dynamics*, *53*(4), 343 – 367, doi:10.1007/s10236-003-0036-9.

Fälthammar, C.-G. (1968), Radial Diffusion by Violation of the Third Adiabatic Invariant, in *Earth's Particles and Fields*, edited by B. M. McCormac, pp. 157–+.

Fei, Y., A. A. Chan, S. R. Elkington, and M. J. Wiltberger (2006), Radial diffusion and MHD particle

simulations of relativistic electron transport by ULF waves in the September 1998 storm, *Journal of Geophysical Research (Space Physics)*, *111*, 12,209–+, doi:10.1029/2005JA011211.

Friedel, R., S. Bourdarie, and T. Cayton (2005), Inter-calibration of magnetospheric energetic electron data, *Space Weather*, *3*, S09B04, doi: 10.1029/2005SW000153.

Friedel, R. H. W., G. D. Reeves, and T. Obara (2002), Relativistic electron dynamics in the inner magnetosphere - a review, *Journal of Atmospheric and Terrestrial Physics*, *64*, 265–282.

Green, J. C., and M. G. Kivelson (2004), Relativistic electrons in the outer radiation belt: Differentiating between acceleration mechanisms, *Journal of Geophysical Research (Space Physics)*, *109*, 3213–+, doi: 10.1029/2003JA010153.

Hilmer, R., G. Ginet, and T. Cayton (2000), Enhancement of equatorial energetic electron fluxes near $L = 4.2$ as a result of high speed solar wind streams, *J. Geophys. Res.*, *105*, 23,311–23,322.

Horne, R. B., and R. M. Thorne (1998), Potential waves for relativistic electron scattering and stochastic acceleration during magnetic storms, *Geophys. Res. Lett.*, *25*, 3011.

Kalman, R. E. (1960), New approach to linear filtering and prediction problems, *American Society of Mechanical Engineers – Transactions – Journal of Basic Engineering Series D*, *82*(1), 35 – 45.

Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge.

Kim, H.-J., and A. A. Chan (1997), Fully adiabatic changes in storm time relativistic electron fluxes, *J. Geophys. Res.*, *102*, 22,107–22,116, doi: 10.1029/97JA01814.

Koller, J., R. H. W. Friedel, and G. D. Reeves (2005), Radiation Belt Data Assimilation and Parameter Estimation, *LANL Reports*, *LA-UR-05-6700*, http://library.lanl.gov/cgi-bin/getfile?LA-UR-05-6700.pdf

Lainiotis, D. G. (1971), Optimal adaptive estimation: structure and parameter adaptation., *IEEE transactions on automatic control*, *AC-16*(2), 160 – 170.

Ljung, L. (1979), Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems., *IEEE transactions on automatic control*, *AC-24*(1), 36 – 50.

Maybeck, P. (1979), *Stochastic Models, Estimation, and Control - Volume 1*, Academic Press, Orlando, Florida.

Meredith, N., R. B. Horne, R. Iles, R. M. Thorne, D. Heynderickx, and R. A. Anderson (2002), Outer zone relativistic electron acceleration associated with substorm-enhanced whistler mode chorus, *J. Geophys. Res.*, *107*, 1144–1144, doi: 10.1029/2001JA900146.

Mitchell, H. L., P. L. Houtekamer, and G. Pellerin (2002), Ensemble size, balance, and model-error representation in an ensemble kalman filter., *Monthly weather review*, *130*(11), 2791 – 808.

Naehr, S., and F. Toffoletto (2005), Radiation belt data assimilation with an extended kalman filter, *Space Weather*, *3*(6), S06,001, doi:10.1029/2004SW000121.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1986), *Numerical Recipes - The Art of Scientific Computing*, 1st ed., Cambridge Univ. Press, Cambridge, United Kingdom.

Reeves, G. D., K. L. McAdams, R. H. W. Friedel, and T. P. O'Brien (2003), Acceleration and loss of relativistic electrons during geomagnetic storms, *Geophys. Res. Lett.*, *30*, 1529.

Reeves, G. D., et al. (1998), The global response of relativistic radiation belt electrons to the January 1997 magnetic cloud, *Geophys. Res. Lett.*, *25*, 3265–3268.

Reeves, G. D., et al. (2005), Toward understanding radiation belt dynamics, nuclear explosion-produced artificial belts, and active radiation belt remediation: Producing a radiation belt data assimilation model, in *Global Physics of the Coupled Inner Magnetosphere*, edited by J. L. Burch, Geophys. Monogr. Ser., pp. –, AGU, Washington, D.C.

Roederer, J. G. (1970), *Dynamics of Geomagnetically Trapped Radiation*, Springer-Verlag, New York.

Russell, C. T., R. C. Snare, J. D. Means, D. Pierce, D. Dearbourne, M. Larson, G. Barr, and G. Le (1995), The GGS/Polar magnetic field investigation, *Space Sci. Rev.*, *71*, 563–582.

Schulz, M., and L. J. Lanzerotti (1973), *Particle Diffusion in the Radiation Belts*, 1st ed., Springer-Verlag, New York.

Schulz, M., and L. J. Lanzerotti (1974), *Particle Diffusion in the Radiation Belts*, Springer-Verlag, New York.

Shprits, Y. Y., R. M. Thorne, G. D. Reeves, and R. Friedel (2005), Radial diffusion modeling with empirical lifetimes: comparison with crres observations, *Annales Geophysicae*, *23*, 1467–1471.

Shprits, Y. Y., R. M. Thorne, R. B. Horne, S. A. Glauert, M. Cartwright, C. T. Russell, D. N. Baker,

and S. G. Kanekal (2006), Acceleration mechanism responsible for the formation of the new radiation belt during the 2003 Halloween solar storm, *Geophys. Res. Lett.,* , *33*, 5104–+, doi:10.1029/2005GL024256.

Sorenson, H. (Ed.) (1985), *Kalman Filtering: Theory and Application*, IEEE Press, New York.

Summers, D. (2005), Quasi-linear diffusion coefficients for field-aligned electromagnetic waves with applications to the magnetosphere, *Journal of Geophysical Research (Space Physics)*, *110*, 8213–+, doi: 10.1029/2005JA011159.

Talagrand, O. (1997), Assimilation of observations, an introduction, *Journal of the Meteorological Society of Japan*, *75*(1B), 191 – 209.

Tarantola, A. (1987), *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier Science Publisher, Amsterdam.

Tarantola, A., and B. Valette (1982), Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion, *Reviews of Geophysics and Space Physics*, *20*, 219–232.

Thomsen, M., D. McComas, G. Reeves, and L. Weiss (1996), An observational test of the Tsyganenko (T89a) model of the magnetospheric field, *J. Geophys. Res.*, *101*, 24,827–24,836.

Tsyganenko, M., H. Singer, and J. Kasper (2003), Storm-time distortion of the inner magnetosphere: how servere can it get?, *J. Geophys. Res.*, *108*, 1209, doi:10.1029/2002JA009808.

Tsyganenko, N. A. (2002), A model of the near magnetosphere with a dawn-dusk asymmetry - 1. mathematical structure, *Journal of Geophysical Research*, *107*(A8), 1179.

Varotsou, A., D. Boscher, S. Bourdarie, R. B. Horne, S. A. Glauert, and N. P. Meredith (2005), Simulation of the outer radiation belt electrons near geosynchronous orbit including both radial diffusion and resonant interaction with Whistler-mode chorus waves, *Geophys. Res. Lett.,* , *32*, 19,106–+, doi: 10.1029/2005GL023282.

J. Koller, Space Science and Applications, Los Alamos National Laboratory, P.O. Box 1663, MS D466, Los Alamos, NM 87545, USA. (jkoller@lanl.gov)

This preprint was prepared with AGU's LaTeX macros v5.01, with the extension package 'AGU^{++}' by P. W. Daly, version 1.6b from 1999/08/19.