1. INTRODUCTION

The highly energetic electron environment in the inner magnetosphere (geosynchronous orbit and inward) has received a lot of research attention in recent years to better understand the dynamics of relativistic electron acceleration, loss, and transport. Physical processes in the Earth’s 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 1970’s that radial diffusion is a key process influencing radiation belt dynamics (Hilmer et al. 2000; Brautigam and Albert 2000). Recently, new observations and increased monitoring (for a review see Friedel et al. (2002)) evidenced that other processes play an important role as well. 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. Despite uncertainties in the precise nature of all the processes controlling radiation belt dynamics, it is widely believed that radial diffusion is one of the critical factors that need to be accurately specified. Boscher et al. (1996) and Bourdarie et al. (1996) show that radial diffusion accounts for 80% of the dynamics for MeV electrons.

Our new approach is to extend available techniques of data assimilation that are widely used for other geophysical systems (meteorology, oceanography, ionosphere) to the radiation belt. The general purpose of data assimilation is to combine measurements and models to produce best estimates of current and future conditions. The resulting “assimilated state” is, either closer to the data or the model depending on their uncertainties. Correlations and uncertainties are incorporated and carried along automatically. The output of the data assimilation is based on all measurements and the model.

One important method of data assimilation is the Kalman filter (Kalman 1960). It became popular because it is a recursive solution to the optimal estimator problem. However, the Kalman filter is just one way of finding an optimized solution. The least square method, for instance, can be extended to provide forecast predictions as well (Tarantola 1987) and the two methods become basically equivalent.

One of the first attempts was the “direct data insertion” technique used by Bourdarie et al. (2005). They showed that by adding data of just one extra satellite into their simulations the updated model could achieve global fidelity on the order of the input data uncertainty - in effect overcoming the fundamental limitations of the underlying physics model. Bourdarie et al. (2005) also described that one of the most crucial, and often overlooked, requirements is the fidelity of satellite data intercalibration. However, obtaining a well calibrated and inter-calibrated set of radiation belt particle data can be a very time consuming but an essential task (Friedel et al. 2005).

While diffusion is an important part in 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 (EMIC) waves, and other plasma waves 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 Lab. It will develop a next-generation space radiation model using extensive satellite measurements, new theoretical insights, global physics-based magnetospheric models, and the techniques of data assimilation. This paper presents the first results from the radiation belt module and also lays the foundation for the data assimilation in the DREAM model.

2. THE MODEL AND DATA FRAMEWORK

2.1 Radial Diffusion Model

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

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

We neglect any source or loss terms here. They are simply additive and we will argue below that these are included implicitly by the data assimilation algorithm.

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 $D_{LL}(Kp, L) = 10^{(0.506Kp-9.325)}L^{10}$ (Brautigam and Albert 2000).

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 of unity. 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 average at unity.

The inner boundary at $L^* = 1$ is fixed at zero however the outer boundary is a free parameter that can be adjusted by the Kalman Filter. We use a so-called augmented state vector approach where the state vector is extended by parameters of the physics model, here the outer boundary.

### 2.2 Ensemble Kalman Filter

The term “data assimilation” is short for model-based assimilation of observations, i.e. data assimilation is the combination of a 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. 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) 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).

The Kalman filter involves three steps that are summarized as follows:

$$\begin{align*}
&\{y^o(t_i), x^i(t_i)\} \\
&\xrightarrow{K} x^a(t_i) \\
&\xrightarrow{M} x(t_{i+1})
\end{align*} \quad (2)$$

where $y^o(t)$ is the observational state vector and $x$ is the state vector. The other terms are as described below.

1. Gain computation: which yields the “Kalman gain matrix” or “weight matrix” $K$.

2. State estimate: which uses the Kalman gain $K$ to weight the “observational residual” (in the older meteorological literature) or the “innovation vector” $d = y^o - Hx$ and computes the “state estimate” or “assimilated state” $\hat{x} = x + K \cdot d$. The operator $H$ is the observational operator which maps the state vector into the observational domain.

3. State forecast or prediction: The next step is to apply a “forward model operator” $M$ which results in the “forecast state vector” $x(t_{i+1})$ that can be compared with new observations at time $t_{i+1}$ in the next cycle. See also Koller et al. (2005).

We use a variant of the classical Kalman Filter since the classical version works only for linear models. We want to leave the outer boundary as a free parameter but doing so makes our model operator non-linear. The ensemble Kalman Filter (Evensen 1994, 2003) can describe the error statistics from non-linear models by using a Monte-Carlo technique. The ensemble members are created by randomly perturbing the state vector, separately advancing them in time 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 ensemble members, after the model has been applied, determines the uncertainty of the forecast. The more ensemble members that are used, the better the probability distribution of the state vector is determined.

### 2.3 Kalman Innovation as Source in the Model

The second step in the Kalman filter warrants a more detailed discussion because this is where we argue that “missing physics” e.g. sources and losses are included implicitly in the Kalman filter method. $x^i$ is calculated with the diffusion equation (1) without the additional source/loss. However, $K \cdot d$ represents effectively a source/loss but depending on the observations and not a physical model. It is important to note here that the magnitude of $K \cdot d$ depends not only on the observations but also 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. On the other hand, if the uncertainty of the model is large, then more weight will be given to the observations.

### 2.4 Data, Model, and Parameter Uncertainties

Uncertainties of the observations $\Delta_y$ and the model $\Delta_M$ are key ingredients for every data assimilation. The ob-
observational uncertainty can be estimated by comparing different satellite measurements of the same parameter against one another and adding an estimate of systematic uncertainties. By using a one dimensional grid with the other dimensions projected onto it, one can find many conjunctions between the geosynchronous satellites. A statistical analysis of the conjunctions will give the relative uncertainty of the observations assuming all instruments have the same uncertainty. We find a relative uncertainty of 30% using 6500 conjunctions. We use the same data uncertainty for GPS and POLAR. We note that in practice this is often only a best estimate of the observational uncertainty.

Model uncertainties, $\Delta_M$, are determined by a combination of the ensemble spread, $\Delta_\epsilon$, 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 do not know the magnitude of the resulting model uncertainty. This is an ongoing research topic in the atmospheric data assimilation community. See (Mitchell et al. 2002) and references therein.

We did several tests and find that an ensemble spread $\Delta_\epsilon = \Delta_\gamma$ leaves enough room for the Kalman Filter to adjust for the fast changes in phase space densities of the observations. The uncertainty of the outer boundary is estimated by how fast the observations from the POLAR satellite change.

3. DATA

We used data from three LANL geosynchronous satellites (LANL-97a, 1991-080, 1990-095), POLAR, and GPS-ns41 for our data assimilation of a six month period in 2002 (Figure 2) top panel.

We obtained the phase space densities at given adiabatic invariants ($\mu = 2083 \text{ MeV}/G, K = 0.1 \sqrt{\text{GR}_E}$) by using the angular resolved electron fluxes and local magnetic field magnitude for each of the satellites. We also applied the global magnetic field configuration from the Tsyganenko 2001 storm 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.

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 ($\mu, J, L^*$).

The POLAR satellite, with a polar orbit of $2 \times 9R_E$, 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 measurement at the apogee equatorial crossings. POLAR also carries a Magnetic Field Experiment (MFE) (Russell et al. 1995) measuring magnetic field vectors. Therefore, $\mu$ can be calculated directly but $J$ and $L^*$ still require the model.

The GPS satellites have a circular orbit with a radius of $4R_E$ 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 T01s model for ($\mu, J, L^*$).

We made strong efforts to calibrate the measured data between satellites. The inter-calibration between the three LANL GEO satellites was obtained by matching the phase space densities (Chen et al. 2005), that is, comparing the phase space density values of electrons with the same combination of ($\mu, K, L^*$) but measured by satellites at different spatial locations during magnetically quiet times. The same method is applied to obtain...
FIG. 2: Observations, reanalysis results, and $Dst$ from July 15 to December 30, 2002. The top panel shows the phase space density data from POLAR, three LANL Geosynchronous satellites, and one GPS satellite. The middle panel depicts the results from our reanalysis with data assimilation by combining observations with a physical diffusion model. White circles mark the L-shell locations where satellite data was assimilated. The bottom panel shows $Dst$ for the same time period.
the inter-calibration between LANL GEO and POLAR (Chen et al. 2006). Also, a preliminary inter-calibration between POLAR and LANL GPS fluxes was done by Friedel et al. (2005).

One distinguished feature in Figure 2 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 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.

4. RESULTS

The result from our reanalysis, the assimilated state, is shown in the middle panel of Figure 2. All regions where no data was available are now filled by the data assimilation procedure as are temporal gaps in the data coverage by individual satellites. However, confidence in the estimated state decreases with distance from location of available data. The white dots in 2 denote the $L^*$ locations of assimilated data.

As with previous studies (Green and Kivelson 2004; Taylor et al. 2004; Chen et al. 2005), we find that the phase space densities have maxima around geosynchronous orbit. All major storms in this time period (see also bottom panel for $Dst$) show a drop-out in phase space density over a large range down to $L = 4$. However, each recovery phase can be very different: While most of them show an enhancement at geosynchronous orbits (e.g. August 21, October 14) some have an enhancement over a much larger $L$-shell range like from $L = 4 – 6$ for September 8, October 25, November 3. There are also two cases (July 22, August 10) where $Dst$ drops only to $Dst = -35$ but the enhancements in the recovery phase are very large and over a wide range of $L = 4 – 8$. This confirms in phase space densities what was seen in fluxes by Reeves (1998).

The first fifteen days of the assimilated state are likely to be more inaccurate because the initial condition for the assimilation was a pre-existing steady state system with a high boundary. It took the ensemble Kalman filter a while to adjust the outer boundary to lower values consistent with the observations (see Figure 1). The outer boundary then stayed low for most of the time. The values for the outer boundary is driven mostly by POLAR data due to the proximity of the computational boundary at $L = 10$.

None of the large accelerations at geosynchronous are explained by radial inward diffusion. The ensemble Kalman Filter did not raise the boundary condition to facilitate the inward diffusion but rather added phase space densities locally like a source term.

The region inside of $L = 4$ is mostly due to the chosen steady state initial condition and some slow inward diffusion modulated by $Kp$. It is not driven by any data and therefore does not show large changes. The correlation between the outer radiation belt and the inner radiation belt is weak.

5. DISCUSSION AND CONCLUSION

We studied the combination of a 1-D radial diffusion code with an ensemble Kalman filter and assimilated data from 5 satellites in the second half of 2002. The data from three LANL geosynchronous, POLAR, and GPS show strong enhancements and drop outs. But the data coverage is only limited. The advantage of reanalysis with data assimilation lies in obtaining a complete picture of the radiation belt. Data gaps are filled by the model and free parameters like the outer boundary can be estimated.

Our ensemble Kalman Filter estimated the state and the outer boundary as a parameter. We find that the outer boundary stayed low most of the time after an initial adjustment period. This result points to the explanation that the phase space density enhancements in the recovery phases are due to local acceleration processes and not radial inward diffusion.

We find that the ensemble Kalman Filter can compensate for a missing source or loss terms in the physics model. This makes the ensemble Kalman Filter an efficient tool to study radiation belt data and holds promise to extract new physical understanding by quantifying how much the Kalman filter had to compensate for. This will be studied in a forthcoming paper.

A reanalysis of historic data can be also very useful for a possible AE/AP-8 replacement. Data assimilation provides the techniques for filling data gaps and carrying the uncertainties along for proper statistics. We are
planning to create a reanalysis of a whole solar cycle that can be used to fly artificial satellites for pre-flight risk assessment. This is a currently ongoing project funded by NASA LWS program in collaboration with Aerospace Corporation.

REFERENCES


