

# An empirical model of magnetospheric chorus amplitude using solar wind and geomagnetic indices

D. I. Golden,<sup>1</sup> M. Spasojevic,<sup>1</sup> W. Li,<sup>2</sup> and Y. Nishimura<sup>2</sup>

Received 14 August 2012; revised 15 October 2012; accepted 16 October 2012; published 7 December 2012.

[1] Magnetospheric chorus is an Extremely Low Frequency/Very Low Frequency (ELF/VLF, 0.3–30 kHz) electromagnetic wave phenomenon which plays an important role in the acceleration and loss of energetic electrons in the Earth's radiation belts. One must therefore possess accurate estimates of chorus amplitudes in order to model radiation belt dynamics. The goal of this study is to design an empirical model of chorus amplitude, the output of which can be used as input to models of radiation belt dynamics. In pursuit of this goal, we compare two related empirical models of chorus amplitude that we have developed based on THEMIS data from June 2008 through December 2011 which use multiple regression to predict equatorial chorus amplitudes as a function of  $L$  and MLT. One model uses only  $AE^*$  and  $K_p$  as model inputs, and the other model utilizes solar wind measurements and geomagnetic indices. The models perform similarly, with each one achieving a median RMS prediction error of  $0.39 \log_{10}$  pT (a factor of 2.5 in amplitude). The coefficients of determination of chorus amplitude for the full model and the  $AE^*/K_p$  model are 0.034 and 0.026, respectively, meaning that these models explain 3.4 and 2.6 percent of the variance of chorus amplitude. We present a parametric analysis, showing the expected effects on chorus amplitude from a modeled substorm and solar wind pressure pulse, as well as modeled chorus amplitude over the course of the month of September 2008. The model outputs give important insight into the global evolution of equatorial chorus amplitude as a function of geomagnetic storm and substorm phase.

**Citation:** Golden, D. I., M. Spasojevic, W. Li, and Y. Nishimura (2012), An empirical model of magnetospheric chorus amplitude using solar wind and geomagnetic indices, *J. Geophys. Res.*, 117, A12204, doi:10.1029/2012JA018210.

## 1. Introduction

[2] Magnetospheric chorus is a naturally occurring Extremely Low Frequency/Very Low Frequency (ELF/VLF, 0.3–30 kHz) electromagnetic wave which is observed within the Earth's magnetosphere and outside the plasmapause. Cyclotron resonant interactions between chorus and energetic electrons in the Earth's radiation belts result in large-scale electron acceleration [e.g., Meredith *et al.*, 2002; Horne *et al.*, 2003, 2005a, 2005b] and loss [e.g., Lorentzen *et al.*, 2001; O'Brien *et al.*, 2003; Thorne *et al.*, 2005; Shprits *et al.*, 2006] in the outer radiation belt. In addition, chorus may be a significant source of plasmaspheric hiss [Bortnik *et al.*, 2008, 2009], which is largely responsible for the slot region between the inner and outer radiation belts [e.g., Lyons and Thorne, 1973]. As a consequence, chorus waves play a significant role in the maintenance of the electron radiation belts [Thorne,

2010]. Typically, timescales for radiation belt electron dynamics are determined via quasi-linear diffusion modeling [e.g., Khazanov *et al.*, 2003; Glauert and Horne, 2005; Albert, 2005]; these models depend on accurate estimates of chorus amplitude which are used as inputs.

[3] In this study, we investigate two potential global models of magnetospheric chorus amplitude: one which uses only instantaneous  $AE^*$  and  $K_p$  as inputs, and one which includes instantaneous and time histories of various solar wind measurements and geomagnetic indices. The presented modeling methodology builds on a similar empirical model of plasmaspheric hiss amplitude, also using THEMIS data [Golden *et al.*, 2012]. Using the output of these models, we show, for the first time, the global evolution of chorus amplitude as a function of solar wind and geomagnetic driving conditions. These models may be used to provide chorus amplitude input to models of global electron diffusion under various geomagnetic driving conditions.

## 2. Instrumentation and Methods

### 2.1. THEMIS Data

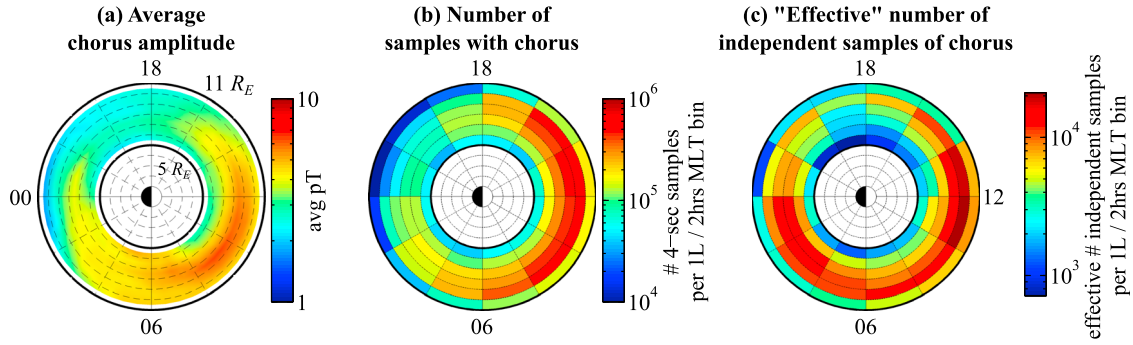
[4] The THEMIS mission [Angelopoulos, 2008] is composed of five probes in near-equatorial orbits with perigee and apogee between 2 and 10  $R_E$ . This study uses survey-

<sup>1</sup>Electrical Engineering Department, Stanford University, Stanford, California, USA.

<sup>2</sup>Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, California, USA.

Corresponding author: D. I. Golden, Electrical Engineering Department, Stanford University, Stanford, CA 94305, USA. (dgolden1@stanford.edu)

©2012. American Geophysical Union. All Rights Reserved.  
0148-0227/12/2012JA018210



**Figure 1.** Chorus statistics derived from THEMIS satellite data for this study period, from 1 June 2008 to 20 December 2011. (a) Average chorus amplitude (the average is taken over the logarithm of chorus amplitude). (b) Number of 4-sec samples where chorus was detected above the instrument noise floor (bin size is  $\Delta L = 1$ ,  $\Delta \text{MLT} = 2$  hrs). (c) “Effective” number of *independent* samples.

mode Search Coil Magnetometer (SCM) [Roux *et al.*, 2008; Le Contel *et al.*, 2008] data from the Digital Fields Board (DFB) [Cully *et al.*, 2008] instrument from June 2008 through December 2011. Magnetic field data were chosen instead of electric field data to avoid contamination by electrostatic (e.g., electron cyclotron harmonic) waves in the measurements. Only THEMIS data which are recorded outside the plasmasphere and within the magnetosphere on the three inner-magnetospheric THEMIS probes (A, D and E) are used in this study. We determine the plasmaspheric and magnetospheric boundaries via the approach described in section 2 of Li *et al.* [2010a], which involves estimating the electron density based on spacecraft potential data from the Electric Fields Instrument [Bonnell *et al.*, 2008] and electron thermal speed measurements from the Electrostatic Analyzer [McFadden *et al.*, 2008].

[5] Chorus amplitude is determined using the upper three filter bank (FBK) channels of the DFB, which have frequency ranges of 80–227 Hz, 316–904 Hz and 1390–4000 Hz (passbank levels 4, 2 and 0, respectively, from Table 6 of Cully *et al.* [2008]). We define the net observed chorus power to be the sum of the power of any of the three FBK channels which are at least partially covered by the typical chorus frequency range of  $0.1\text{--}0.7 f_{ceq}$  [Burtis and Helliwell, 1976], where  $f_{ceq}$  is the equatorial gyrofrequency along the given field line, computed using the dipole model of the Earth’s magnetic field. This requirement that a channel need only partially lie in the range  $0.1\text{--}0.7 f_{ceq}$  is based on the assumption that chorus is the only measured wave with significant  $B$ -field amplitude in this frequency range and no wave power will lie outside the range  $0.1\text{--}0.7 f_{ceq}$ . If any fraction of the frequency range  $0.1\text{--}0.7 f_{ceq}$  is not covered by these three FBK channels, the data are discarded in order to avoid missing chorus power that is in the range  $0.1\text{--}0.7 f_{ceq}$  but is outside of the measured frequency range. Based on these criteria, all FBK data where THEMIS is in the range  $L < 5.3$  ( $f_{ceq} > 5.7$  kHz) or  $L > 10.3$  ( $f_{ceq} < 800$  Hz) are discarded because the chorus frequency range cannot be adequately covered by the upper three FBK channels. We also exclude instrument anomalies, which appear as very high or very low channel amplitude. Because the FBK data set contains data from only one of the three SCM antennas, and the measured  $B$ -field component of this antenna is perpendicular to the spacecraft spin axis, the measured

chorus amplitude is somewhat below the true amplitude; if, on average, the angle between the antenna measurement and the  $B$ -field is  $45^\circ$ , then the measured wave amplitude will be below the true amplitude by, on average, a factor of  $\sqrt{2}$ . We have not corrected for this effect in the chorus amplitude presented in this manuscript.

[6] As in Golden *et al.* [2012], we bin THEMIS chorus measurements by dipole  $L$  and MLT into bins with  $\Delta L = 1$  and  $\Delta \text{MLT} = 2$  hrs. Figure 1a shows average observed chorus amplitude as a function of  $L$  and MLT. Amplitude values appear “smoothed” because they are linearly interpolated between bin centers. Figure 1b shows the number of 4-sec observations of chorus per bin. Chorus appears with maximum average amplitude in the pre-noon sector ( $06 < \text{MLT} < 14$ ), though the MLT region with the greatest number of observations of chorus (at any detectable amplitude) extends somewhat later to  $\text{MLT} = 18$ .

[7] Samples of chorus amplitude are drawn from a time series, and as a result, successive samples are dependent on each other (i.e., the chorus amplitude at time  $t + 1$  can be predicted fairly accurately from the amplitude at time  $t$ ). Figure 1c shows the “effective” number of *independent* samples in each bin [Bayley and Hammersley, 1946] assuming that measurements of chorus amplitude are a first order autoregressive, i.e., AR(1), process. Under the assumption of an AR(1) process, the effective number of independent samples for calculations of the sample mean is calculated as

$$n' = n \left\{ 1 + \frac{2}{n} \frac{1}{1-a} \left[ a \left( n - \frac{1}{1-a} \right) - a^n \left( 1 - \frac{1}{1-a} \right) \right] \right\}^{-1} \quad (1)$$

from Mudelsee [2010, equation 2.7], where  $n$  is total number of samples and  $a$  is the autocorrelation coefficient. The effective number of samples per bin is used to avoid overfitting the model, as described in the next section.

## 2.2. Modeling Methodology

[8] As in the model of hiss from Golden *et al.* [2012], we model chorus amplitude from solar wind and geomagnetic indices using a linear multiple regression model with time histories of  $SYM-H$ ,  $AE$ ,  $P_{\text{dyn}}$ ,  $d\Phi_{MP}/dt$  and  $P_{\text{dyn}}^{1/2} \Phi_{MP}/dt$ . Each  $L/\text{MLT}$  bin is modeled independently. In each bin, we consider the possibility of including the current value of each solar wind measurement and geomagnetic index (with

respect to the time of the THEMIS chorus measurement) as well as the average value between 0 and 0.25 hrs ago, 0.25 and 0.5 hrs ago, etc., up to the average between 8 and 16 hrs ago (the width of the moving average window doubles with each step after the first; there are 8 steps total, including the instantaneous value). This logarithmic spacing of the averaging window is in contrast to the linear spacing from Golden *et al.* [2012] which extended up to 6 hrs of history. We hypothesized in this study that we could apply coarser averages for more distant data history, which would allow us to include more data history (up to 16 hrs) without dramatically increasing the number of potential features in our model (and thereby increasing the computational burden).

[9] We use sequential feature selection with 10-fold cross validation to choose relevant input features for the model in each bin. Sequential feature selection is an automated procedure which proceeds as follows. Beginning with a model with only a constant term, each measurement time history (a “candidate feature”) is added to the model. The augmented model is subjected to 10-fold cross validation based on 10 separate partitions of the data. Cross-validation determines whether a model that includes that candidate feature and is specified on 90% of the data (nine of the partitions) is better or worse at predicting the remaining 10% of the data (remaining partition), based on RMS error, when compared to a model which does not include that candidate feature. The validation step is run 10 times on each subset of the data, with each data point appearing once in the 10% test set, and the RMS error is averaged over the 10 iterations. The candidate feature which best improves the cross-validation performance of the model is added to the model, and the procedure begins again with the remaining candidate features. If at any point no additional candidate feature improves the performance of the model, then sequential feature selection is complete. To reduce the potential for overfitting, we also terminate the sequential feature selection procedure before the number of features exceeds 1/10 of the effective number of independent samples of chorus in that bin (Figure 1c).

[10] Though it is theoretically possible for the model in a given bin to have no features selected (besides a constant), in practice, all bins were specified with models that had at least one non-constant feature. All bins were specified independently, and were not constrained to have the same selected model features as their neighbors. Because the chorus amplitude data are autocorrelated, to reduce the similarity between the 10 different cross validation partitions, the partitions are specified contiguously; e.g., the first partition contains the earliest recorded 10% of data, the second partition contains the next-earliest recorded 10% of data, etc.

[11] For each bin, we used least squares to calculate coefficients for the linear model, using the input features that were chosen via sequential feature selection. These coefficients are supplied in the auxiliary material.<sup>1</sup>

[12] Besides this model with solar wind measurements and geomagnetic indices as features (the “full model”), we specified an additional model for comparison which uses only  $AE^*$  (the maximum value of the  $AE$  index in the

preceding three hours) and  $K_p$  as features (the “ $AE^*/K_p$  model”). This model was also specified with different coefficients in each bin using least squares, but because the features (only  $AE^*$  and  $K_p$ ) were common to all bins, sequential feature selection was not necessary.

### 3. Goodness of Fit

[13] We employ two separate metrics to assess the “goodness of fit” for both the full (solar wind and geomagnetic index) model and the  $AE^*/K_p$  model, shown in Figure 2. The first metric is the coefficient of determination (the square of the correlation coefficient,  $r$ ) between the true chorus amplitude and the chorus amplitude predicted by the model (higher numbers are better). The coefficient of determination may be interpreted as “the proportion of the variance of measured chorus amplitude explained by the model” [e.g., Chatterjee and Hadi, 2008, section 7.2]. A perfect model will have  $r^2 = 1$ , while a model whose output is completely uncorrelated with the true chorus amplitude (e.g., a model which always predicts a constant amplitude) will have  $r^2 = 0$ .

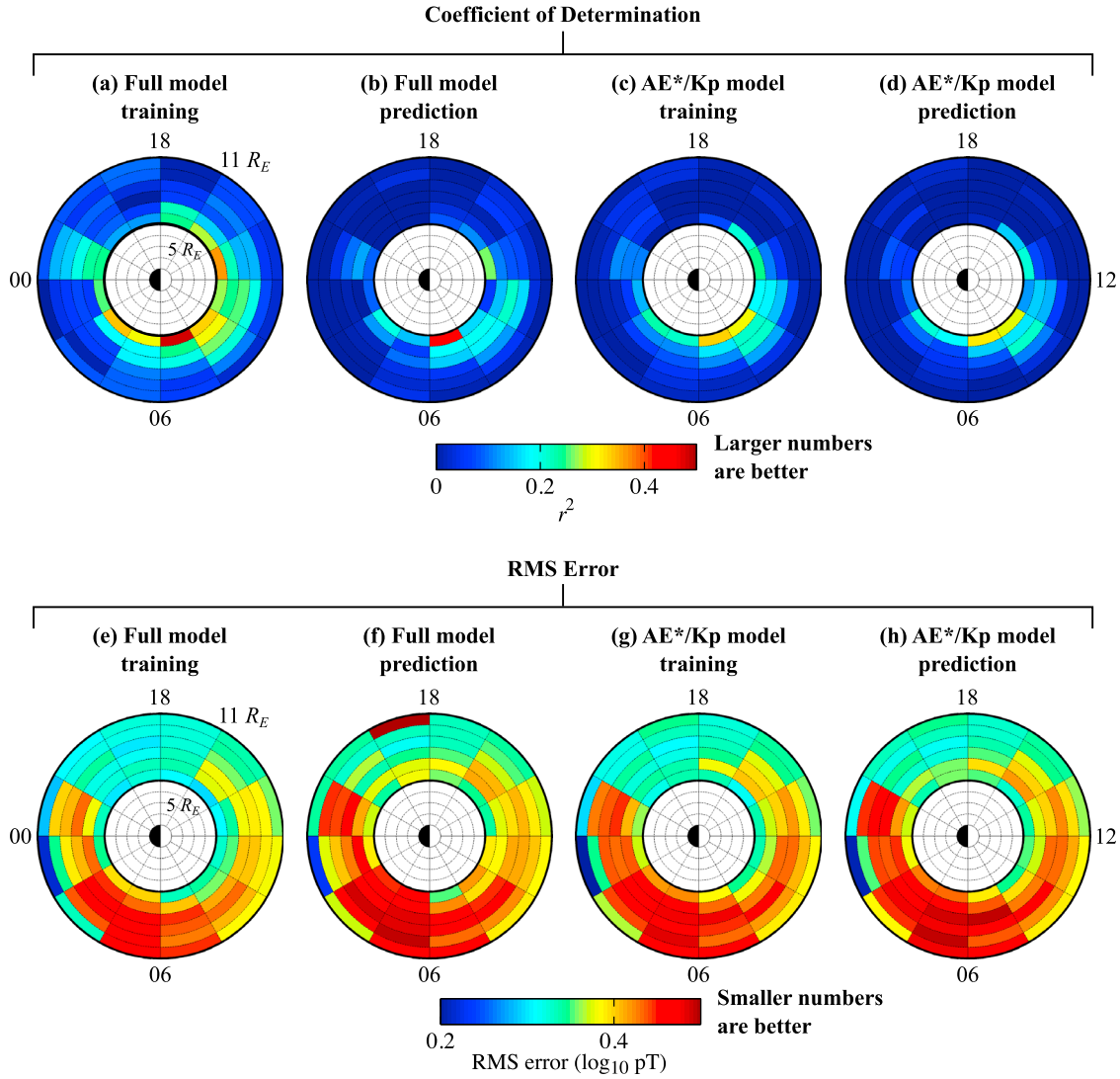
[14] The coefficient of determination is not a perfect metric, because in regions with uniformly low chorus amplitude (e.g., the pre-midnight sector in Figure 1a), one may not be as concerned with how well the predicted amplitude tracks the true amplitude as long as the predicted amplitude is also uniformly low. Therefore, we also use the root mean square (RMS) error between the true and the predicted chorus amplitude as a metric of model accuracy (lower numbers indicate better fit). The advantage of this metric is that it gives an explicit measure of the accuracy of the model, in units of wave amplitude (or relative amplitude). The disadvantage is that the RMS error will decrease as the variance of chorus amplitude decreases, e.g., on the dusk side where the average chorus amplitude is low (Figure 1a), regardless of the quality of the model; for example, a null model with no features except for a constant, which always predicts the global mean of chorus amplitude, will have lower RMS error when predicting values with low variance than when predicting values with high variance.

[15] Figure 2 shows both the model training and prediction error. Training error is defined as the model performance when tested on the same data used to specify the model, while prediction error is the model performance when tested on new data. Training error is determined by specifying a model in each bin (following the procedure of section 2.2) and then testing the resulting model on the same data. To determine prediction error, we first partition the data into 10 segments, each of which is contiguous in time. Then, for each segment, the model is specified on the other nine segments and tested on the remaining segment. The goodness-of-fit metric is then averaged over all 10 iterations.

[16] The coefficient of determination for the full model’s training set tends to be higher (better) than that for the prediction set; this suggests that this model is moderately overfit to the training set. The full model’s RMS training error is also very slightly lower than its prediction error, but the difference is more subtle. The  $AE^*/K_p$  model performance shows no evidence of overfitting.

[17] Prediction performance for both the full model and the  $AE^*/K_p$  model are similar, both showing the best

<sup>1</sup>Auxiliary material data sets are available at <ftp://ftp.agu.org/apend/ja/2012ja018210>. Other auxiliary material files are in the HTML. doi:10.1029/2012JA018210.



**Figure 2.** Model training and prediction error for full model which uses solar wind data and geomagnetic indices as inputs and model which uses only current  $AE^*$  and  $K_p$  as inputs. (a–d) The coefficient of determination (square of the correlation coefficient,  $r$ ) between the true and modeled chorus amplitude. (e–h) The root-mean square error between the true and modeled chorus amplitude.

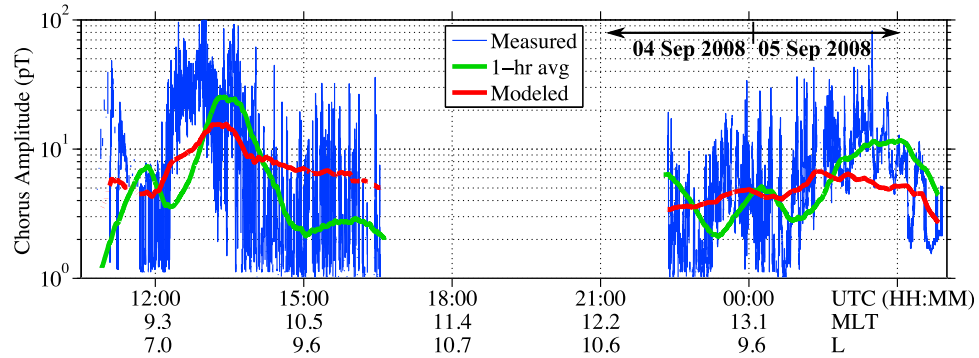
performance, as measured by the coefficient of determination, in the pre-noon sector ( $06 < \text{MLT} < 12$ ) in approximately the region where the average chorus amplitude is highest (Figure 1a). Both model's RMS prediction error is lowest where the chorus amplitude is lowest on the dusk side ( $14 \lesssim \text{MLT} \lesssim 22$ ), which is largely due to the fact that the variance of chorus amplitude is lowest in that region. The median coefficient of determination is slightly more favorable for the full model than for the  $AE^*/K_p$  model, but the median RMS error is approximately the same (Table 1). The median is used instead of the mean because the per-bin distribution of errors is not normal.

[18] To gain greater insight into sources of error in the full model, we simulated a satellite pass using the model during the time of an actual THEMIS satellite pass (Figure 3). This simulated pass shows that the modeled chorus amplitude changes much more slowly than the measured chorus

amplitude; the modeled chorus amplitude appears to vary on a similar timescale to the 1-hr running average of the measured chorus amplitude, though the absolute amplitudes are different. The smoothed appearance of the chorus model output is likely due to the manner in which we included time histories as features in the model, which involve averages over up to 8 hours of data. Since these features are running averages, they force a smoothing of the model output in time. Thus, we might not expect a model which includes

**Table 1.** Model Performance

	Full Model	$AE^*/K_p$ Model
Median coefficient of determination prediction (unitless, higher is better)	0.034	0.026
Median RMS error prediction ( $\log_{10}$ pT, lower is better)	0.39	0.39



**Figure 3.** THEMIS A chorus amplitude on 04–05 Sep 2008 along with the 1-hr running average of the measured chorus amplitude and the modeled chorus amplitude over the same path. Modeled amplitude is linearly interpolated from model bin centers to the satellite location.

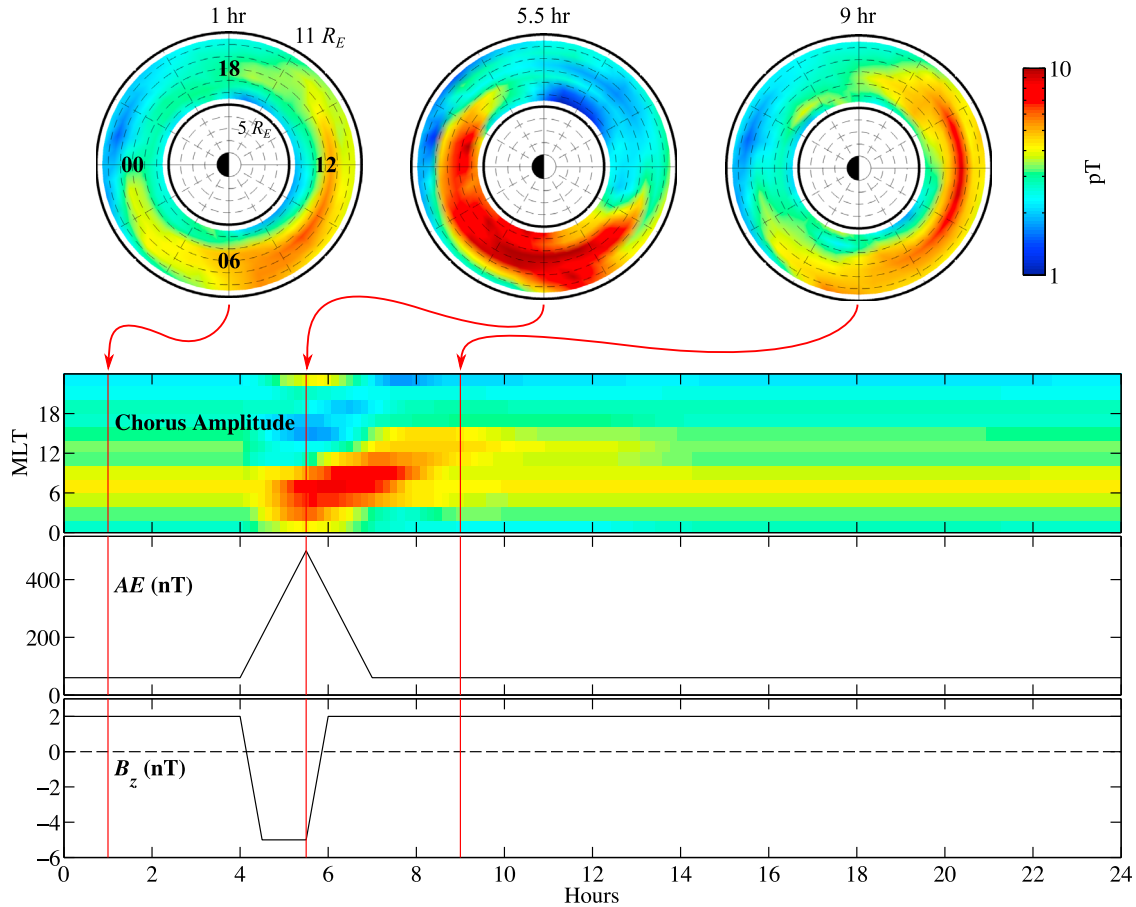
these moving average features to be able to track the short timescale variation in chorus amplitude.

#### 4. Example Model Output

[19] We investigated the global response to a geomagnetic substorm by simulating substorm conditions in the model inputs and examining the model output. We simulated two related substorms: one involving a triangular “pulse” in the

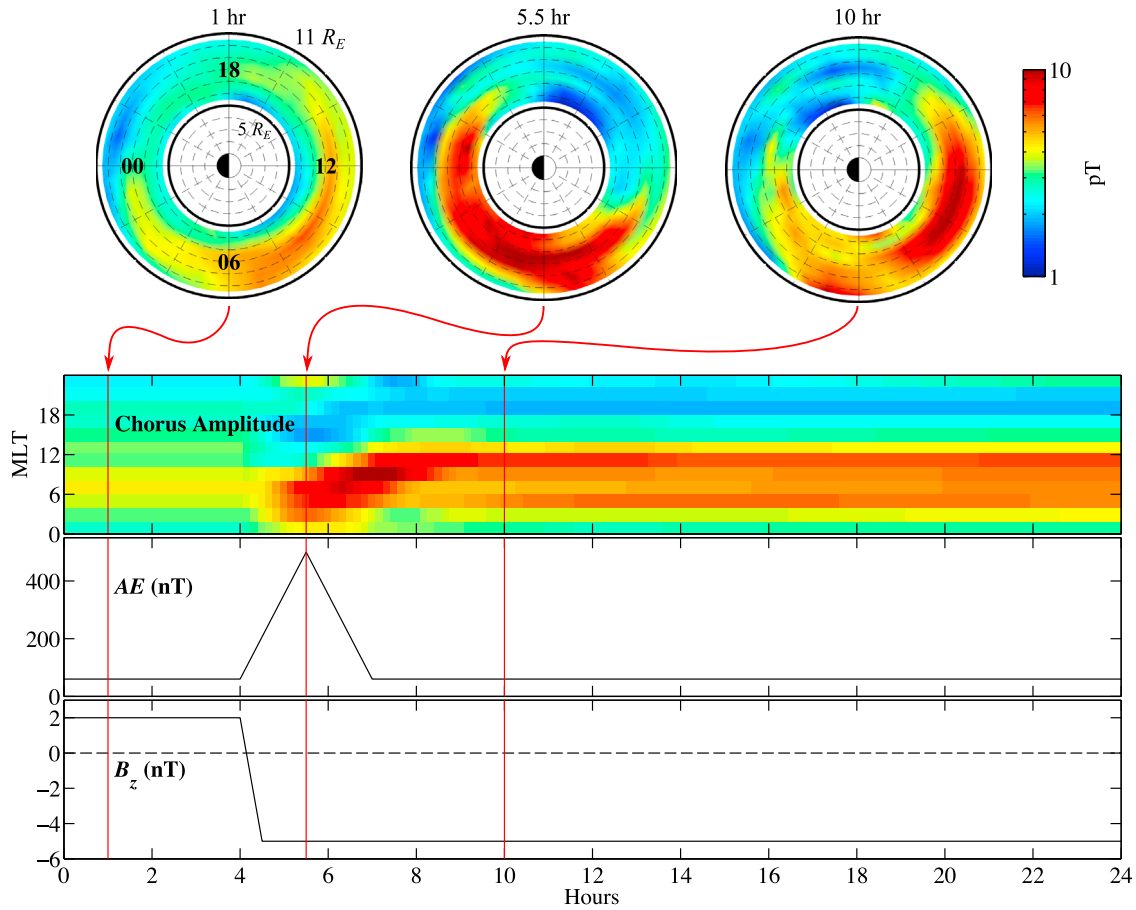
$AE$  index and a square pulse in  $B_z$  turning from northward to southward (Figure 4) and one involving the same triangular  $AE$  pulse with a “step” in  $B_z$  turning from northward to southward and remaining southward for the remainder of the simulation (Figure 5). The remaining model features ( $B_y$ ,  $P_{\text{dyn}}$ ,  $v_{\text{sw}}$  and  $SYM-H$ ) were held at their approximate median values.

[20] In both cases, the chorus amplitude initially rapidly increases on the dawn side from  $22 < \text{MLT} < 10$ . The high-



**Figure 4.** Output of full model in response to a simulated substorm with a two-hour negative  $B_z$  pulse. Chorus amplitude versus MLT and time is averaged over  $5 \leq L \leq 11$ .  $AE$  and  $B_z$  were varied as shown while the remaining model inputs were held at  $B_y = 0$  nT,  $P_{\text{dyn}} = 1.5$  nPa,  $v_{\text{sw}} = 375$  km/sec,  $SYM-H = -5$  nT.





**Figure 5.** Same as Figure 4 but  $B_z$  is varied as a step instead of a pulse.

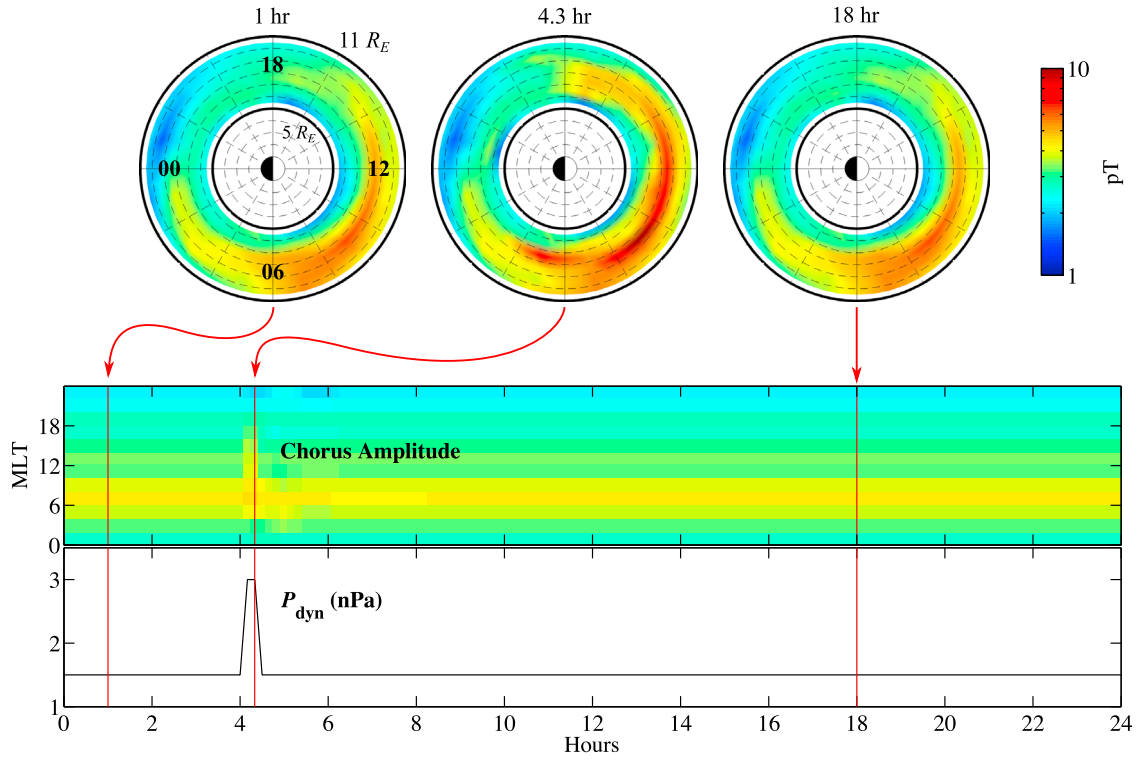
amplitude region rotates eastward from the dawn side to the day side from  $04 < \text{MLT} < 14$ . In the case of the  $B_z$  pulse (Figure 4), the chorus amplitude is lowered and returns to ambient conditions within  $\sim 4$  hours of the end of the pulse, while for the  $B_z$  step (Figure 5), the chorus amplitude remains elevated for the remainder of the simulation, despite the fact that  $AE$  returns to its ambient level. Chorus also appears to be suppressed in the post noon sector ( $12 < \text{MLT} < 18$ ) near the peak in  $AE$  in both simulations, and it is suppressed in the pre-midnight sector ( $18 < \text{MLT} < 00$ ) following the  $B_z$  step. A movie of the simulated substorm with the  $B_z$  pulse appears in the auxiliary material.

[21] The rate of rotation of the chorus amplitude peak in MLT is approximately 3 hrs MLT per hr UTC, giving a 24-hr MLT drift period of 8 hrs. Since chorus is generated by azimuthally drifting energetic electrons, we expect that the rate of rotation of the chorus peak should be equal to the drift rate of the seed electron population. For energetic, equatorially trapped electrons, an 8-hour drift period at  $L \sim 7$  corresponds to  $\sim 10$ – $20$  keV electrons [Walt, 1994, Figure B.1]. This energy range is slightly lower than the typically quoted range for chorus generation of  $>40$  keV [Tsurutani and Smith, 1974], but consistent with the minimum resonant energy of a few to tens of keV as recently calculated based on measured plasma parameters on THEMIS [Li et al., 2010b,

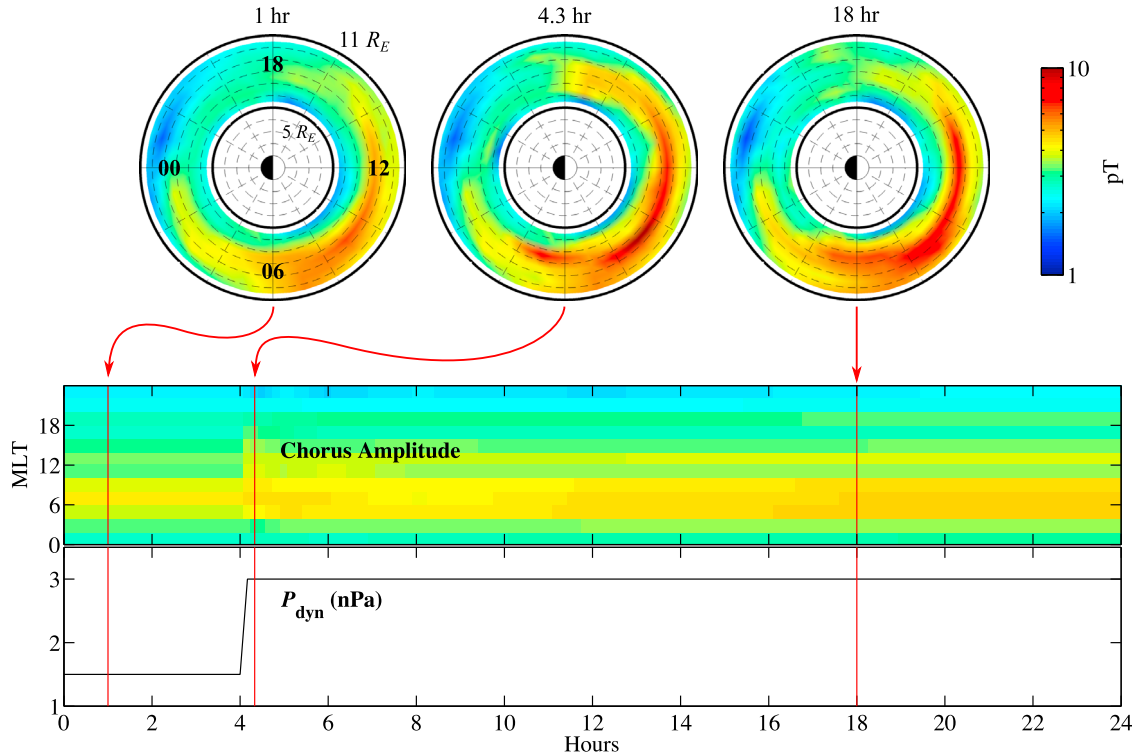
section 3.2]. The suppression of waves in the post-noon sector is possibly the effect of the global convection electric field opening the drift paths of the source electrons such that the electrons drift toward the magnetopause rather than into the dusk sector.

[22] We also investigated the effect of a moderate pressure pulse (Figure 6) and pressure step (Figure 7) via the full chorus model. Solar wind pressure pulses and intervals of elevated pressure have been shown observationally to be a driver of chorus particularly in the dayside outer magnetosphere [Gail et al., 1990; Salvati et al., 2000; Spasojevic and Inan, 2010]. From a theoretical standpoint, solar wind compression of the dayside magnetosphere leads to  $L$ -shell splitting of the source electrons drifting from dawn toward noon [e.g., Roederer, 1967]. Drift shell splitting results in anisotropic electron distributions at high  $L$  in the noon sector, which are conducive for chorus generation. This effect is believed to be at least partially responsible for the high overall occurrence rate of chorus in the outer dayside [Li et al., 2009; Spasojevic and Inan, 2010; Bunch et al., 2012]. Enhanced solar wind dynamic pressure will amplify the effect of drift shell splitting.

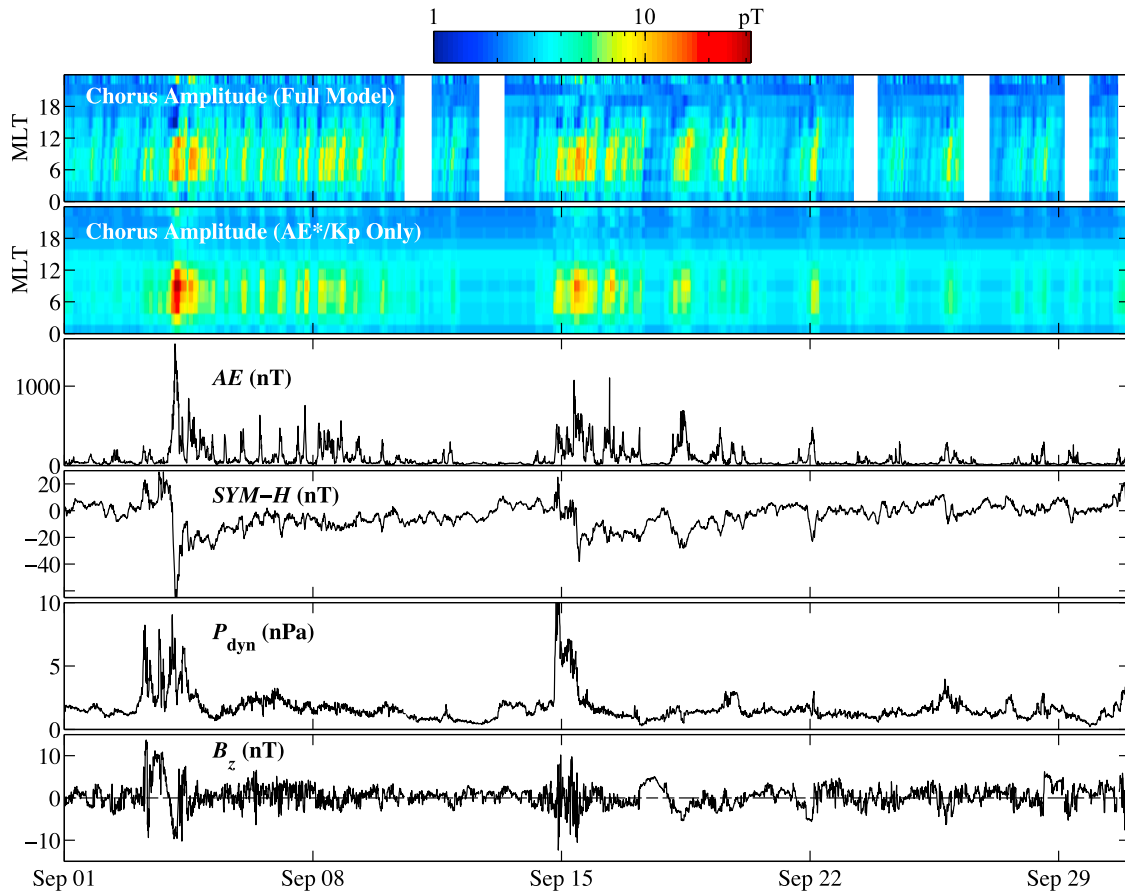
[23] In practice, both the simulated pressure pulse and the simulated step result in very modest increases in chorus amplitude. In the case of the pulse, the response does not last much longer than the pulse itself. In the case of



**Figure 6.** Output of full model in response to a simulated pulse of increased dynamic pressure. Chorus amplitude versus MLT and time is averaged over  $5 \leq L \leq 11$ .  $P_{\text{dyn}}$  was varied as shown while the remaining model inputs were held at  $B_z = 2$  nT,  $B_y = 0$  nT,  $v_{\text{sw}} = 375$  km/sec,  $AE = 60$  nT,  $SYM-H = -5$  nT.



**Figure 7.** Same as Figure 6 but  $P_{\text{dyn}}$  is varied as a step instead of a pulse.



**Figure 8.** Output of full model and  $AE^*/K_p$ -only model of chorus amplitude for the month of September 2008, along with relevant geomagnetic indices and solar wind parameters. Modeled chorus amplitude is shown versus MLT and time, averaged over  $5 \leq L \leq 11$ . Both chorus models show increases in chorus amplitude as a result of substorms (small peaks in  $AE$ ) and storms (4 September, 15 September). Only the full chorus model captures the dynamic eastward rotation of the chorus peak amplitude, which appears as upward sloped chorus peaks in the top panel during storms and substorms. Periods when the full model could not be computed completely due to lack of solar wind data availability are shown in white.

the step, the chorus amplitude steadily increases over the course of the simulation as pressure remains high although chorus amplitude never increases by much more than 1 pT over the ambient level.

[24] Finally, we also ran the full model and the  $AE^*/K_p$  model using real data from the entire month of September 2008. This month included several dozen substorms as well as a few minor storms, including one on 4 September with min  $SYM-H = -67$  nT and one on 15 September with min  $SYM-H = -39$  nT. The output of both models is shown in Figure 8. The outputs of the two models are very similar, as might be expected since their prediction errors are so similar (Figure 2). During storms and substorms (as defined by, e.g., peaks in  $AE$ ), chorus amplitude is briefly amplified and then returns to a background level. Only the full chorus model captures the dynamic Eastward rotation of the chorus region, which appears as upward-sloped peaks (movement to later MLTs versus time) in the top panel during storms and substorms. The 24-hr MLT drift period of the chorus in the full model output is approximately 6.5 hrs, with minimal variation between the different substorms. This drift period corresponds to an equatorial electron energy of  $\sim 20$  keV.

A movie of the full model run over this period appears in the auxiliary material.

## 5. Discussion

[25] The outputs of the presented chorus models may be used as inputs to radiation belt models to provide more accurate estimates of in situ chorus amplitude than are currently used. Before beginning this study, we had hypothesized that the full model, which uses time histories of geomagnetic indices and solar wind parameters, would be more accurate (e.g., would have significantly lower median RMS error) than a model which included only  $AE^*$  and  $K_p$ . Such a discrepancy is seen when modeling plasmaspheric hiss via similar methods [Golden et al., 2012]. It was therefore surprising to see that the full model and the  $AE^*/K_p$  models have nearly the same RMS error and therefore are equally effective at predicting in situ chorus amplitude. This similarity in model performance is likely due to the fact that both models vary their outputs very slowly in response to changing input conditions (Figure 3) because both are constructed of features with long time constants; the full model contains moving average terms of up to



8 hours,  $AE^*$  is a 3-hour “moving maximum” term and  $K_p$  is a 3-hour index. Therefore, modeled chorus amplitude can be interpreted as a time average of the instantaneous amplitude, over a timescale of up to 3 (for the  $AE^*/K_p$  model) or 8 hours (for the full model). The RMS error may be improved in future models by not including any moving average (or similar) terms, and including only terms with short (e.g., 1-min) time constants. At the moment, however, the presented full and  $AE^*/K_p$  models may be used interchangeably to generate inputs for models of radiation belt dynamics.

[26] One advantage of the full model over the  $AE^*/K_p$  model is its ability to give insight into the time-behavior of chorus due to the inclusion of time histories of its features. The full model was used to show how, on average, the chorus region tends to rotate Eastward during storms and substorms, a behavior which is not apparent from the output of the  $AE^*/K_p$  model (Figure 8). The rotation period of the chorus is consistent with a source equatorial electron population of 10–20 keV.

## 6. Conclusion

[27] We have presented two novel empirical models for global estimates of equatorial chorus amplitude. One model uses  $AE^*$  and  $K_p$  as inputs, and the other model uses solar wind measurements and geomagnetic indices as inputs. In terms of mean squared prediction error, the two models perform very similarly, making them both adequate for generating estimates of chorus amplitude to use as inputs to radiation belt models. However, the full solar wind and geomagnetic index model is more appropriate for qualitatively evaluating the dynamics of the in situ chorus environment. The shorter timescale features included in the full model allow it to represent certain dynamic features, such as the eastward rotation of the chorus source region during storms and substorms, that are not well represented by the  $AE^*/K_p$  model.

[28] **Acknowledgments.** The work at Stanford University was supported by the NSF under awards 1043442 and 1137411 and by NASA under award NNX09AF51G. The work at UCLA was supported by NASA under awards NNX11AR64G and NNX11AD75G.  $AE$  and  $SYM-H$  data are from the World Data Center for Geomagnetism, Kyoto. Solar wind data are from the Virtual Radiation Belt Observatory, based on data from OMNIWeb, courtesy of R. Denton and Z. Qin, who acknowledge support from NSF award 0751007. THEMIS data are from NASA CDAWeb, and we acknowledge NASA contract NAS5-02099 and V. Angelopoulos for use of THEMIS data, J. W. Bonnell and F. S. Mozer for EFi data, C. W. Carlson and J. P. McFadden for ESA data, A. Roux and O. LeContel for SCM data and C. Cully for FBK data.

[29] Robert Lysak thanks the reviewers for their assistance in evaluating this paper.

## References

- Albert, J. M. (2005), Evaluation of quasi-linear diffusion coefficients for whistler mode waves in a plasma with arbitrary density ratio, *J. Geophys. Res.*, **110**, A03218, doi:10.1029/2004JA010844.
- Angelopoulos, V. (2008), The THEMIS mission, *Space Sci. Rev.*, **141**, 5–34, doi:10.1007/s11214-008-9336-1.
- Bayley, G. V., and J. M. Hammersley (1946), The “effective” number of independent observations in an autocorrelated time series, *J. R. Stat. Soc. Suppl.*, **8**(2), 184–197.
- Bonnell, J. W., F. S. Mozer, G. T. Delory, A. J. Hull, R. E. Ergun, C. M. Cully, V. Angelopoulos, and P. R. Harvey (2008), The Electric Field Instrument (EFI) for THEMIS, *Space Sci. Rev.*, **141**, 303–341, doi:10.1007/s11214-008-9469-2.
- Bortnik, J., R. M. Thorne, and N. P. Meredith (2008), The unexpected origin of plasmaspheric hiss from discrete chorus emissions, *Nature*, **452**, 62–66, doi:10.1038/nature06741.
- Bortnik, J., W. Li, R. M. Thorne, V. Angelopoulos, C. Cully, J. Bonnell, O. LeContel, and A. Roux (2009), An observation linking the origin of plasmaspheric hiss to discrete chorus emissions, *Science*, **324**, 775–778, doi:10.1126/science.1171273.
- Bunch, N. L., M. Spasojevic, and Y. Y. Shprits (2012), Off-equatorial chorus occurrence and wave amplitude distributions as observed by the Polar Plasma Wave Instrument, *J. Geophys. Res.*, **117**, A04205, doi:10.1029/2011JA017228.
- Burtis, W. J., and R. A. Helliwell (1976), Magnetospheric chorus: Occurrence patterns and normalized frequency, *Planet. Space Sci.*, **24**, 1007–1007, doi:10.1016/0032-0633(76)90119-7.
- Chatterjee, S., and A. S. Hadi (2008), *Regression Analysis by Example*, 2nd ed., Wiley-Interscience, Hoboken, N. J.
- Cully, C. M., R. E. Ergun, K. Stevens, A. Nammari, and J. Westfall (2008), The THEMIS Digital Fields Board, *Space Sci. Rev.*, **141**, 343–355, doi:10.1007/s11214-008-9417-1.
- Gail, W. B., U. S. Inan, R. A. Helliwell, D. L. Carpenter, and S. Krishnaswamy (1990), Characteristics of wave-particle interactions during sudden commencements: 1. Ground-based observations, *J. Geophys. Res.*, **95**, 119–137, doi:10.1029/JA095iA01p00119.
- Glauert, S. A., and R. B. Horne (2005), Calculation of pitch angle and energy diffusion coefficients with the PADIE code, *J. Geophys. Res.*, **110**, A04206, doi:10.1029/2004JA010851.
- Golden, D. I., M. Spasojevic, W. Li, and Y. Nishimura (2012), Statistical modeling of plasmaspheric hiss amplitude using solar wind measurements and geomagnetic indices, *Geophys. Res. Lett.*, **39**, L06103, doi:10.1029/2012GL051185.
- Horne, R. B., S. A. Glauert, and R. M. Thorne (2003), Resonant diffusion of radiation belt electrons by whistler-mode chorus, *Geophys. Res. Lett.*, **30**(9), 1493, doi:10.1029/2003GL016963.
- Horne, R. B., R. M. Thorne, S. A. Glauert, J. M. Albert, N. P. Meredith, and R. R. Anderson (2005a), Timescale for radiation belt electron acceleration by whistler mode chorus waves, *J. Geophys. Res.*, **110**, A03225, doi:10.1029/2004JA010811.
- Horne, R. B., et al. (2005b), Wave acceleration of electrons in the Van Allen radiation belts, *Nature*, **437**, 227–230, doi:10.1038/nature03939.
- Khazanov, G. V., K. V. Gamayunov, and V. K. Jordanova (2003), Self-consistent model of magnetospheric ring current and electromagnetic ion cyclotron waves: The 2–7 May 1998 storm, *J. Geophys. Res.*, **108**(A12), 1419, doi:10.1029/2003JA009856.
- LeContel, O., et al. (2008), First results of the THEMIS search coil magnetometers, *Space Sci. Rev.*, **141**, 509–534, doi:10.1007/s11214-008-9371-y.
- Li, W., R. M. Thorne, V. Angelopoulos, J. Bortnik, C. M. Cully, B. Ni, O. LeContel, A. Roux, U. Auster, and W. Magnes (2009), Global distribution of whistler-mode chorus waves observed on the THEMIS spacecraft, *Geophys. Res. Lett.*, **36**, L09104, doi:10.1029/2009GL037595.
- Li, W., R. M. Thorne, J. Bortnik, Y. Nishimura, V. Angelopoulos, L. Chen, J. P. McFadden, and J. W. Bonnell (2010a), Global distributions of suprathermal electrons observed on THEMIS and potential mechanisms for access into the plasmasphere, *J. Geophys. Res.*, **115**, A00J10, doi:10.1029/2010JA015687.
- Li, W., et al. (2010b), THEMIS analysis of observed equatorial electron distributions responsible for the chorus excitation, *J. Geophys. Res.*, **115**, A00F11, doi:10.1029/2009JA014845.
- Lorentzen, K. R., J. B. Blake, U. S. Inan, and J. Bortnik (2001), Observations of relativistic electron microbursts in association with VLF chorus, *J. Geophys. Res.*, **106**, 6017–6028, doi:10.1029/2000JA003018.
- Lyons, L. R., and R. M. Thorne (1973), Equilibrium structure of radiation belt electrons, *J. Geophys. Res.*, **78**, 2142–2149, doi:10.1029/JA078i013p02142.
- McFadden, J. P., C. W. Carlson, D. Larson, M. Ludlam, R. Abiad, B. Elliott, P. Turin, M. Marckwardt, and V. Angelopoulos (2008), The THEMIS ESA plasma instrument and in-flight calibration, *Space Sci. Rev.*, **141**, 277–302, doi:10.1007/s11214-008-9440-2.
- Meredith, N. P., R. B. Horne, D. Summers, R. M. Thorne, R. H. A. Iles, D. Heynderickx, and R. R. Anderson (2002), Evidence for acceleration of outer zone electrons to relativistic energies by whistler mode chorus, *Ann. Geophys.*, **20**, 967–979.
- Mudelsee, M. (2010), *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*, *Atmos. Oceanic Sci. Libr.*, vol. 42, Springer, Dordrecht, Netherlands, doi:10.1007/978-90-481-9482-7.
- O’Brien, T. P., K. R. Lorentzen, I. R. Mann, N. P. Meredith, J. B. Blake, J. F. Fennell, M. D. Looper, D. K. Milling, and R. R. Anderson (2003), Energization of relativistic electrons in the presence of ULF power and MeV microbursts: Evidence for dual ULF and VLF acceleration, *J. Geophys. Res.*, **108**(A8), 1329, doi:10.1029/2002JA009784.

- Roederer, J. G. (1967), On the adiabatic motion of energetic particles in a model magnetosphere, *J. Geophys. Res.*, *72*, 981–992, doi:10.1029/JZ072i003p00981.
- Roux, A., O. Le Contel, C. Coillot, A. Bouabdellah, B. de La Porte, D. Alison, S. Ruocco, and M. C. Vassal (2008), The Search Coil Magnetometer for THEMIS, *Space Sci. Rev.*, *141*, 265–275, doi:10.1007/s11214-008-9455-8.
- Salvati, M. A., U. S. Inan, T. J. Rosenberg, and A. T. Weatherwax (2000), Solar wind control of polar chorus, *Geophys. Res. Lett.*, *27*, 649–656, doi:10.1029/1999GL010702.
- 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*, L05104, doi:10.1029/2005GL024256.
- Spasojevic, M., and U. S. Inan (2010), Drivers of chorus in the outer day-side magnetosphere, *J. Geophys. Res.*, *115*, A00F09, doi:10.1029/2009JA014452.
- Thorne, R. M. (2010), Radiation belt dynamics: The importance of wave-particle interactions, *Geophys. Res. Lett.*, *37*, L22107, doi:10.1029/2010GL044990.
- Thorne, R. M., T. P. O'Brien, Y. Y. Shprits, D. Summers, and R. B. Horne (2005), Timescale for MeV electron microburst loss during geomagnetic storms, *J. Geophys. Res.*, *110*, A09202, doi:10.1029/2004JA010882.
- Tsurutani, B. T., and E. J. Smith (1974), Postmidnight chorus: A substorm phenomenon, *J. Geophys. Res.*, *79*, 118–127, doi:10.1029/JA079i001p00118.
- Walt, M. (1994), *Introduction to Geomagnetically Trapped Radiation*, Cambridge Univ. Press, Cambridge, U. K.