Predicting lower band chorus with autoregressive-moving average transfer function (ARMAX) models

Laura E. Simms[1], Mark J. Engebretson[1], Craig J. Rodger[2], Jesper W. Gjerloev[3], and Geoffrey D. Reeves[4]

[1] Department of Physics, Augsburg University, Minneapolis, MN, USA

[2]Department of Physics, University of Otago, Dunedin, New Zealand

[3] Johns Hopkins University, Laurel, MD, USA

[4] Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Key Points:

1. Empirically based autoregressive-moving average transfer function models predict daily-averaged lower band chorus power well
2. Validation correlation between chorus observations and predictions using source electron flux and substorm activity as inputs was 0.675
3. Source electron flux showed an immediate influence on chorus while substorms were influential over many days

Abstract

We model lower band chorus observations from the DEMETER satellite using daily and hourly autoregressive-moving average transfer function (ARMAX) equations. ARMAX models can account for serial autocorrelation between observations that are measured close together in time and can be used to predict a response variable based on its past behavior without the need for recent data. Unstable distributions of radiation belt source electrons (tens of keV) and the substorm activity (SMEd from the SuperMAG array) that is thought to inject these electrons were both statistically significant explanatory variables in a daily ARMAX model describing chorus. Predictions from this model correlated well with observations in a hold-out test data set (validation correlation of 0.675). Source electron flux was most influential when observations came from the same day or the day before the chorus measurement, with effects decaying rapidly over time. Substorms were more influential when they occurred on previous days, presumably due to their injecting source electrons from the plasma sheet. A daily ARMAX model with IMF |B|, IMF Bz, and solar wind pressure as inputs instead of those given above was somewhat less predictive of chorus (r=0.611). An hourly ARMAX model with only solar wind and IMF inputs was even less successful, with a validation correlation of 0.502.

1. Introduction

VLF chorus waves (very low frequency: 0.3-10 kHz discrete waves) play an important role in transferring energy in the magnetosphere. They are thought to be partially responsible for accelerating electrons to high energies (Bortnik & Thorne, 2007; Li et al., 2005; Meredith et al., 2002; Simms et al., 2018a; Summers et al., 2007; Thorne et al., 2013; Xiao et al., 2015) as well as precipitating electrons into the ionosphere through pitch-angle scattering into the loss cone (Bortnik et al., 2006; Bortnik & Thorne, 2007; Hendry et al., 2012; Hikishima et al., 2010; Horne and Thorne, 2003; Lam et al., 2010; Lorentzen et al., 2001; Millan and Thorne 2007; Orlova and Shprits, 2010). Lower band chorus waves (0.1-0.5 of the fce (electron cyclotron frequency)) may be the most effective at driving electron acceleration (Horne and Thorne 1998; Meredith et al., 2003; Summers et al., 1998). For this reason, it is important to reach an understanding of what magnetospheric processes drive lower band chorus waves and to develop models to predict their levels.

Unstable distributions of source electrons (tens of keV) are postulated to drive chorus waves, with these electrons injected from the plasma sheet during substorm activity (Li et al., 2009; Meredith et al., 2001; 2003; Rodger et al., 2016; Su et al., 2014; Tsurutani & Smith, 1974). Substorm activity itself often results from increased activity in the solar wind (Lyons et al., 2005; McPherron et al., 2009).

VLF power spectral density (PSD) observations from the DEMETER satellite provide an opportunity to model chorus activity due to these drivers. In our analyses, we explore several approaches. We use two sets of parameters to produce predictive models: 1) radiation belt source electron flux (31.7 keV) and substorms, and 2) IMF and solar wind parameters. For the measure of substorm activity we use the SMEd (darkside) auroral electrojet index from SuperMAG (Gjerloev, 2012). This nightside index has a clear physical meaning as it is a better indicator of nightside activity (e.g., substorms) than the AE index (Newell and Gjerloev, 2011) and thus has a closer relationship with plasma sheet conditions. We produce models using simple correlations, followed by multiple regression using lagged predictor variables from one hour previous (Lag1) and one day previous (Lag24).

However, correlation and regression do not, on their own, account for possible serial autocorrelation of the response variable (chorus power). In time series data, an observation may be highly correlated with those near to it in time. This can not only invalidate tests of statistical significance, but it may also produce models that show instability in predicting novel observations. In addition, there may be cyclical patterns in the time series that could be accounted for, thereby producing a more accurate model.

One approach to the problem of periodicity is to introduce a collection of sine and cosine terms to the regression model, essentially using Fourier terms to approximate periodic trends in the data (Hyndman & Athanasopoulos, 2018). Seasonality and other periodicity can then be removed from the data by subtraction. For an example of this approach applied to space weather data, see Borovsky and Denton (2014).

However, another approach to time series analysis is the ARIMA model. This uses a collection of autoregressive (AR), moving average (MA), and differencing (I) terms to model the behavior of the response variable. This can be easily extended to add input (or exogenous) variables which describe the influence of other parameters on the response variable. These are called ARIMAX models (with the X coming from the eXogenous variables), dynamic regression models, or transfer function models. In these models, past and current values of the explanatory variables (the inputs) influence the response variable, however, the response variable does not have any effect on the input variables (Makridakis et

al., 1998). Once transfer models are trained, they do not require further input of the response variable to produce accurate predictions.

An ARMAX approach to modelling space weather phenomenon has been used previously to study the nonlinear effects of solar wind drivers of high energy electrons (Balikhin, et al., 2011; Boynton, et al. 2013ab). We use it here to model the response of lower band chorus to source electron flux and substorms, and, subsequently, to a set of IMF and solar wind parameters which may be more readily available as inputs to a predictive model.

2. Data

Satellite-observed VLF lower band chorus power spectral density data (PSD) (log10($\mu V^2/m^2/Hz$)) from 2005 through 2009 were obtained from the ICE (Instrument Champ Electrique) on DEMETER in sun-synchronized orbit (Berthelier et al., 2006). This low-Earth polar orbit satellite (2004-2010) was switched off at high latitudes so most observations occurred in McIlwain L shells 2-4 over roughly ± 45-75 °. Although VLF waves are thought to be produced near the equator, despite the limited latitudinal coverage of DEMETER, observations from this satellite have been used to represent chorus in several studies (Neal et al., 2015; Rodger et al., 2016; Santolik et al., 2006; Simms et al., 2018ab; Zhima et al., 2013). We binned L shells such that, observations at L=2.0 – 2.99, 3.0-3.99, 4.0-4.99, 5.0-5.99, 6.0-6.99 were classified as L= 2, 3, 4, 5, and 6, respectively. Lower band chorus observations (.1 - .5fce) were hourly averaged over the dayside passes of the satellite (1030 LT) and the nightside passes (2230 LT). Initial correlations were performed for all L bins and both dayside and nightside observations. Time series models were then developed for dayside observations at L=4.

For "source" electrons (tens of keV), we use daily averaged 31.7 keV log electron fluxes (log(electrons /(cm^2-s-sr-keV)) measured by the Synchronous Orbit Particle Analyzer (SOPA) instrument on the Los Alamos National Laboratory (LANL) satellites in geosynchronous orbit (Reeves et al., 2011). This channel was previously identified as being less correlated with the somewhat higher "seed" electrons (hundreds of keV) that are accelerated to relativistic energies (Simms et al., 2018a). For the measure of substorm activity we use the SMEd (darkside) auroral electrojet index from SuperMAG (Gjerloev, 2012). Nightside SME has been shown to effectively sample the magnetotail plasma sheet (Newell and Gjerloev, 2011). In addition, hourly averages of IMF Bz component (nT), IMF magnetic field magnitude |B| (nT), Dst (nT), and solar wind velocity (V, in km/sec), proton number density (N in #/cc), and pressure (nPa) are from the OMNIWeb database (collected at L1 orbit and subsequently shifted to expected magnetosphere arrival times). Daily averages of the solar wind and IMF over 24-48 hours previous were calculated from these hourly averages for use in the daily averaged models.

3. Statistical methods

Statistical analyses were performed in IBM SPSS Statistics and MATLAB.

3.1. Basic regression techniques

Coefficients of the multiple regression models are calculated using least squares regression (Neter et al., 1985). These result in linear combinations:

$$\hat{Y} = \beta_0 + \beta_1 \times X_1 + \beta_2 \times X_2 + \cdots \varepsilon \quad (1)$$

where $\hat{Y}$ is the predicted value of the dependent variable (e.g., lower band chorus power), the β are the coefficients describing the effect of each predictor variable (the X's, e.g., possible drivers of chorus), and ε represents the error term, or difference between the predicted ($\hat{Y}_i$) and the observed (Y_i) values for each observation. A valid model would result in random, normally distributed errors with homogeneous variance. A non-random distribution of residuals can alter the statistical significance of the predictors, leading to false conclusions about the influence of parameters. It may also result in poor predictive ability of the model. One form of non-random association of errors is often seen in time series data, where observations close to each other in time may be correlated (Hyndman and Athanasopoulos, 2018).

A model can be produced from a subset of the data (the training set) and then validated with the remaining data (the test set) by calculating the predicted values for the test set from the trained model. The predicted and observed values from the test set can then be correlated to determine how well the model predicts novel data. The validation correlation (*r* between observed and predicted values in the test set) is different from the prediction efficiency (also known as the coefficient of determination, R^2) which is often used as a measure of model goodness of fit. The prediction efficiency is the proportion of variation in the training set data explained by the model, with the remaining variation described by the error term. While reducing the unexplained variation is a goal of most modelling efforts, it is not, by itself, a reliable indicator of how well the model will predict novel data. Validation is a more robust test of this.

When comparing the effects of predictors on a common scale, standardized regression coefficients can be calculated by standardizing the variances of all variables in the model to 1. These coefficients then represent how many standard deviations the dependent variable will change when a particular predictor changes by one standard deviation. We use these standardized coefficients to compare effects. However, unstandardized coefficients are reported for the predictive models to allow for new predictions from observed data.

3.2. Time series analyses

The analysis of time series data presents a special problem in statistical methods as each observation may be serially autocorrelated with previous observations. If this is not accounted for, tests of statistical significance may be inflated and give a false picture of the true relationship between variables. Predictions from a model showing serial autocorrelation issues may also not be close to observed values. With non-time series data, the goal of a regression model is to describe the relationship between dependent and predictor variables such that the errors from the model are randomly distributed. If the dependent variable is a time series, however, some of its behavior may be best represented by its own values at previous time periods. Without the addition of terms to describe the effect of past behavior on future values, the errors of the model can be strikingly non-random (which is

a sign of a poorly fitted model) and predictions from the model are therefore suspect. Autocorrelation in a model can be checked by inspecting the partial autocorrelation function (PACF) of the residual errors at each lag.

The simplest time series model, and one which can be easily fit using least squares regression, is the autoregressive model. In this type of model, the current value of the dependent variable is predicted by its past behavior. The shorthand description of this model is AR(p), where p is the number of lags used to describe the observation. The most basic model in the class is a first-order autoregressive model, AR(1), where each observation is dependent on its value in the previous time step plus some random error, ε , in a linear combination with coefficients determined by least squares regression:

$$\hat{Y}_t = \beta_0 + \beta_1 \times Y_{t-1} + \varepsilon \qquad (2)$$

where $\hat{Y}_t$ is the predicted value of the dependent variable at time t. The constant term, β_0, is the overall mean of the time series. Higher order autoregressive models can be used to describe more complicated behavior. For example, an AR(2) model (with both a lag 1 and lag 2 term) could describe a return to the mean as a sinusoidal oscillation.

The simple AR models assume a stationary mean over the time series, but if there is a trend in the data over time this can be accounted for by differencing the observations, i.e., by subtracting a previous value of the response variable. Our data did not show such a trend so we do not address that issue in the models we describe in our current study.

Another way to correct for autocorrelated errors is to regress on a moving average of past values. This is mathematically equivalent to adjusting for the error produced by the last forecast:

$$\hat{Y}_t = \beta_0 + \beta_1 \times \varepsilon_{t-1} \qquad (3)$$

However, due to the fact that the error can only be obtained after the model is fit, the moving average model is no longer able to be fit by the analytical method of least squares.

As a practical matter, the best fit to the data may result from a mix of these approaches. A mixed model is called an ARIMA(*p,d,q*) model: AR for the autoregressive component, I for the differenced component, and MA for the moving average component. The autoregressive order is represented by *p*, the order of differencing by *d*, and the moving average order by *q*. Thus, equation 2 represents ARIMA(1,0,0) (one autoregressive order, no differencing or moving average component), while equation 3 represents ARIMA(0,0,1) (no autoregressive term or differencing, and one order of moving average component). As our data did not show a long term trend, the *d* in all our models is 0.

In addition, "seasonality" may be added to the model to account for repeating patterns that occur over the course of a year (seasons) or, as in this study, a daily pattern. Each of these "seasonal" terms can be AR, differencing, or MA at the cadence of the seasonality. The shorthand description of a model with one order of AR, no differencing, and one order of MA at both the non-seasonal and seasonal level would be ARIMA (1,0,1)(1,0,1) with the numbers in the second set of parentheses representing the AR, difference, and MA terms for the seasonal effect. First-order in the seasonal effect would mean a lag of one "season", e.g., 12 months for a yearly pattern, or 24 hours for a daily pattern.

It is often of interest to include other predictors in the model, either to improve the forecast or to determine their influence. These other predictors can be added as lagged, differenced, or moving

average terms of themselves as well. This type of model is called a transfer function (ARIMAX), as the information from the other predictors is transferred to the dependent variable. It may also be referred to as a dynamic regression model where current and past values of the explanatory variable(s) are presumed to influence the response variable, with the effect of the predictor decreasing exponentially over time (Pankratz 1991):

$$Y_t = a + \frac{\omega(B)}{\delta(B)} X_{t-b} + N_t \quad (4)$$

where $\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \cdots - \omega_s B^s$

$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \cdots - \delta_r B^r$

$N_t = ARIMA\ error\ process$

Here, *a, r, s,* and *b* are constants with *b* being the number of time steps before the predictor variable has an effect on *Y*, and *s* and *r* the number of time steps over which the numerator and denominator act, respectively. *B* in this equation is the backshift operator which is used to simplify the writing of the equation. Essentially, applying this operator moves the variable back in time by the specified number of steps, such that $B^1 X$, for example, is equivalent to X_{t-1}.

Due to the ARIMA nature of this model, current values of *Y* are influenced by its own past values which are in turn influenced by past values of *X*. Thus, the effect of *X* on *Y* at time *t* is a cumulative function of all past values of *X*. Its influence at each lag can be determined by an arithmetic recursive operation which accounts for both the effect at a given lag (the ω coefficients) and the decay of that effect (the exponentially weighted δ coefficients). These models are typically fit using maximum likelihood estimation instead of the simple least squares method if there are moving average terms, and they can be extended to more than one predictor variable (Makridakis et al., 1998).

As with multiple regression models, validation of a time series model can be accomplished by training the model on a subset of the data (the training set) and producing predictions from a withheld subset (the test set). However, as the model depends on past sequential behavior of the dependent variable, it is best to produce the model from a complete time series rather than subsets of data at various disconnected times. Predictions from the model can be produced using the current and past values of the explanatory variables and only the time series behavior of the response variable. Thus, once the model is fit from the training set, further predictions do not require additional input of the response variable. Although a simple forecast model can be made by predicting the response variable using recently observed values, it is often more successful to predict the response based on the previously observed overall behavior rather than the actual values at specific time steps (Hyndman and Athanasopoulos, 2018).

3.3. Model building

We begin with simple correlations between source electron fluxes (31.7 keV) and lower band chorus power at L2-5 to test the hypothesis that variations in the flux of these electrons drive chorus power

changes. We also correlate chorus power with SMEd, which we use as a measure of substorms, as substorms are believed to drive the source electrons which subsequently trigger chorus. Then, in a full regression model, using dayside observations from L4, we test the predictive power of source electron flux and SMEd (on the same day and the previous day) as well as solar wind and IMF parameters (in the same hour, the previous hour, and 24 hours previous). We validated this model by correlating observations with predictions in the test data subset. However, this model showed high correlation of residuals at 1 lag.

In order to reduce the serial autocorrelation, we fit several time series models. One difficulty of time series analysis is that the models work best if there are observations for every time step. Unfortunately, as is common with space weather data, our data set contains periods of missing data. For this reason, we chose a time period where missing data was at a minimum (yearday 2006290-2008174) and where gaps were generally 1-2 hours. We then replaced the missing data with averages of the surrounding hours. However, the data gaps were extensive at L5. Thus, although data from this L shell showed a high correlation between chorus and source electron flux and SMEd, we did not perform time series analysis on this L shell.

As ARIMAX models can become quite complicated with differencing, autoregressive and moving average terms, and predictor variables at a variety of lags and delays, we use the Expert Modeler procedure in SPSS to automatically select a "best" model that is optimized for explaining the variation in the data with the least number of parameters. This procedure searches through the space of possible models and calculates the normalized Bayesian Information Criterion (BIC) for each (Schwarz 1978). A lower BIC means a better fit to the data used to produce the model (i.e., the training set). However, as simply adding more complexity to the model (more terms) may appear to increase the fit, the BIC also penalizes more complex models for increased number of terms. The model with the lowest BIC, therefore, is the one with both the best fit and the least complexity. Both AR and MA terms, as well as transfer function input, can be chosen from any time delay by the automatic procedure.

We fit a daily time series transfer function model (ARMAX) using geostationary source electron flux and SMEd which are presumed to be the most direct drivers of whistler mode chorus activity. We fit daily models because both the flux and SMEd data are essentially at a daily cadence. This ARMAX model was able to predict chorus power variations using current observations of flux and SMEd without the need for new observations of chorus power as an input. For comparison, we fit a daily model using the previous day's chorus as input instead of a predicted AR term, together with the flux and SMEd observations.

We also fit a daily ARMAX model using solar wind and IMF parameters. Although the Expert Modeler procedure had multiple parameters to choose from (solar wind velocity, number density, and pressure, as well as IMF |B| and Bz) some were found to be redundant and hence were not included in the model. Finally, we also fit an hourly transfer function model with IMF and solar wind parameters as the inputs. We tested for "seasonality" in this hourly model with a 24 hour period of seasonality.

For each model, we correlated predictions with the test data set of 2009.

Because none of our data required differencing to correct for a trend, we drop the *i* designation in the following models, referring to them as AR, ARX, and ARMAX.

4. Results

4.1. Source electron flux and substorms as predictors of chorus

A time series plot (Sept 2006, L4 dayside) shows that radiation belt source electron flux and SMEd both tend to track with chorus observations (Figure 1). However, the correspondence is not exact, and there are times when electron flux seems to be the better predictor and other times when SMEd is more predictive. The simple (linear) correlations of daily averaged chorus power with source electron flux (solid lines of Figure 2ab) are higher than with SMEd (solid lines of Figure 2cd) on the current day and 24 hours previous, for each L-shell over L2-5 for both dayside and nightside passes of the satellite. The highest correlation of chorus with source flux (r=0.785 at L4) is higher than that with SMEd (r=0.481), suggesting that the influence of source electrons is higher than that of substorms, even though the substorms are widely presumed to provide the source electrons. For both electron flux and SMEd, the L2 linear correlations are low, while the L3-5 correlations are all higher. Dayside pass chorus power on the same day (i.e., at lag 0) is somewhat more strongly correlated with these two predictors than the nightside observations. Lag 24 predictors (i.e., data from one day previous) have a lower association with chorus. However, these simple correlations are uncorrected for serial autocorrelation.

Simple correlation analysis assumes a linear relationship between variables. To explore if a linear association was the best descriptor of these relationships we also fit a fourth-order polynomial regression model (including linear, quadratic, cubic, and the fourth-order terms) to these single variable models (dashed lines of Figure 2). Fitting a polynomial curve had little effect on the correlation between chorus power and source electron flux. For example, while the linear correlation of chorus with source electron flux at L4 (dayside) was 0.785, the polynomial fit correlation was 0.789. This relationship appears to be linear.

Conversely, at L2, and to some extent at L3, a nonlinear fit gave a higher correlation for the chorus power and SMEd relationship. However, at L4, on the dayside, the polynomial fit between SMEd and chorus was not much better than the linear fit. At L4 the linear correlation of chorus with SMEd was 0.481, while the polynomial fit was 0.529. These were similar enough that we continued fitting linear models for this subset.

For the more highly correlated L4 dayside observations, we produced a regression model using daily averages of chorus power as the response variable and electron flux and SMEd as the predictor variables (Model 1 of Table 1). Unstandardized coefficients are given in the prediction equation (see Data section for units). This would allow prediction of new observations from source electron flux and SMEd in their original scaling:

$$L4\,\widehat{Chorus} = -9.06 + 1.07 \times SourceFlux + 0.618 \times LagSourceFlux + 0.0022 \times SMEd - 0.0018 \times LagSMEd \quad (5)$$

All terms are statistically significant (p-value < 0.05), indicating that the there is only a 5% probability this strong of an influence from each term would be seen by chance alone. For this model, R^2 =0.644,

indicating that the model explains 64.4% of the variation in the training set data. Although the R^2 is sometimes called the prediction efficiency, a validation correlation between observed values in a test set with the predictions from the regression model is a more useful indicator of how well the model forecasts new observations. We predict chorus using this model for the 2009 data (that was withheld as the test set). The validation correlation between observed and predicted values is r= 0.638 (Figure 3a). However, there is also evidence of strong serial autocorrelation. The PACF shows a strong and significant peak at one day, with a partial autocorrelation of 0.310 (Figure 3b). This means we are violating one of the assumptions of linear regression and may have produced a model that inflates the significance levels of the predictors.

To some extent, we can correct the serial autocorrelation in this model by adding the previous day's chorus power observations to the above regression model as an autoregressive term:

$$L4\,\widehat{Chorus} = -5.50 + .909 \times SourceFlux + 0.110 \times LagSourceFlux + 0.003 \times SMEd - 0.003 \times LagSMEd + 0.410 \times LagL4Chorus \quad (6)$$

The prediction efficiency, R^2, is slightly higher, at 0.698 (69.8% of variation explained), and there is a tighter relationship between observed chorus and that predicted by the model (Figure 3c; r=0.766; Model 2 of Table 1). The maximum autocorrelation in the PACF is now 0.103, which is an improvement over the previous model (Figure 3d).

4.2. Regression model with all possible predictors

A daily model including daily averages of all possible parameters (observed chorus, electron flux, SMEd, and both IMF and solar wind variables averaged overthe same day and the day before) resulted in a higher R^2 (0.775) and validation correlation (0.823), and a lower maximum PACF (0.082) (Model 3 of Table 1).

However, predicting hourly chorus power proved to be more difficult. We used source electron flux and SMEd at the same hour (Lag0) and also from the hour 24 hours before (Lag24); solar wind and IMF parameters at Lag0, lagged by 1 hour (Lag1) and Lag24; and observed chorus power from 1 hour before (Model 4 of Table 1). Further lags were not used as they complicated an already cumbersome model. (See the hourly ARMAX model below for a more refined estimation of effects over further lags.) The unstandardized coefficients for the resulting prediction model for dayside L4 chorus are given in Table 2.

All predictors are influential during at least one point in time. However, the validation correlation was low (0.452). Although more information has been added to the model, the erratic hourly behavior was not as well predicted as the smoother daily averages.

Although we report unstandardized coefficients in the above models, we also report the standardized coefficients of Model 4 to show the relative influence of predictors (Figure 4a). Source electron flux has a stronger influence when measured on the same day (Lag0) than when measured on the previous day

(Lag24). It is more influential than SMEd, and SMEd shows a negative effect when measured 24 hours before. Of the other parameters, V has a stronger positive influence on the same day (Lag0), indicating that it may drive processes associated with chorus enhancement that are stronger than the source electron flux effect. Bz shows its strongest effect at Lag1 (one hour before). In other words, a strong (negative) Bz in the hour before is associated with increased chorus. However, including all terms in the model may result in counterintuitive influences. The negative substorm, V, and N effects may indicate that substorms and IMF/solar wind parameters are describing the same phenomena and that their relative influences are confounded with each other. Similarly, the confounding of a parameter's influence with itself at different lags may lead to untrustworthy conclusions.

An R^2 of 0.562 for this hourly model is lower than that of the daily model (above). The PACF shows there is still high autocorrelation between the residual errors of the hourly model even when the previous hour's chorus is added as a predictor (0.453, Figure 4b; Model 4 of Table 1). The validation correlation of this model (r = 0.452, Figure 4c) is lower than that for the daily models. The noisiness of the hourly data is difficult to predict and adding the hourly lag of chorus did not fix the autocorrelation problem.

4.3. Daily ARMAX transfer function models

ARMA models provide a flexible model building technique to deal with serial autocorrelation in time series data. As described above, various combinations of autoregressive (AR) and moving average (MA) terms can be used to precisely model the behavior of the response variable. Once these factors are removed, predictor variables can be tested for their influence on the dependent variable without inflation of significance tests. Transfer function models are designated with an X (i.e., ARX or ARMAX models). In addition to reducing serial autocorrelation, ARMAX models, once trained on known observations, may be able to predict into the future without additional input of the lagged response variable.

We first use source electron flux and SMEd observations as predictors. Chorus power at L4 did not show a trend over time so we did not difference the time series. We allowed the SPSS Expert Modeler procedure to choose AR and MA terms of the chorus power behavior and to pick numerator and denominator terms for the predictor variables. This resulted in an AR(1), MA(4) model with no differencing, or ARMAX(1,0,4). The autoregressive term is estimated from the model output from the day before (*t-1*) while the moving average term is measured from the error term of the model 4 days before (*t-4*). In the numerator coefficients of the transfer function inputs, both predictors act on chorus power on the same day (*t*), with SMEd also acting at a one day lag (*t-1*). The automated procedure did not add any lags previous to this as they did not improve the fit of the model. The effects of both predictors decay over time (denominator coefficients) (Model 5 of Table 3):

$$L4\ \widehat{Chorus}_t = -8.825 + \frac{.831 \times Source_t}{1 - .502 \times Source_{t-1}} + \frac{.004 \times SMEd_t + .004 \times SMEd_{t-1}}{1 - .889 \times SMEd_{t-1}} + .264 \times L4\ Chorus_{t-1} - .120 \times \varepsilon_{t-4} \quad (7)$$

The numerator of the transfer function terms (source electron flux and SMEd) describe when the effect from that parameter begins, while the denominator describes how the effect decays. For example, a one unit increase in log(source electron flux) on the same day results in a 0.831 unit increase in log(chorus power), or that an increase of 10 electrons /(cm^2-s-sr-keV) in source electron flux results in an approximate increase of 6.8 $\mu V^2/m^2/Hz$ in chorus power. The rate at which influence decays is also incorporated into this term via the denominator where the effect of source flux drops off by a factor of 0.502 at each lag. However, the effect of source electron flux and SMEd from previous days on chorus power is even more long-lasting via their prior influence on previous values of chorus power.

The ARMAX model with these two explanatory variables results in a validation correlation of r=0.675 (Figure 5a). This ARMAX model was successful at reducing the serial autocorrelation. The maximum value in the PACF was 0.081 (Figure 5b; Model 5 of Table 3) These predictions are based on the past behavior of daily chorus power measurements and do not include actual measurements of lagged chorus power to calculate predicted power values in the test set. Nevertheless, the model predicts observed power reasonably well (Figure 6c, green line, Model 5). There is little difference in the timeplot between values predicted by the ARMAX model and that of the regression model using source flux, SMEd, and previous day's average chorus power as predictors (blue line of Figure 5d; Model 2). Note, however, that the ARMAX model is able to predict chorus power even when observed chorus is unavailable as an input to the model.

To show the influence of the lagged effects of the inputs (source electron flux and SMEd) over time, we use the standardized regression coefficients to compare the influence of source flux and SMEd on the same scale over a two-week period (Figure 7). While source electron flux and SMEd show a similar effect on the same day, there is a stronger effect of SMEd from the day before. In addition, the lower decay rate of the SMEd effect means that this parameter shows an influence over several weeks while the effect of source electron flux drops off within days.

Current solar wind and IMF inputs are often more readily available from the OMNIWeb data site than either source electron flux or the SMEd index. If more immediate predictions are required, a model using these more available inputs is needed. To produce a daily predictive model using the more obtainable solar wind and IMF inputs, we again used the SPSS Expert Modeler procedure to choose parameters. Using daily averaged solar wind and IMF parameters, the best ARMAX model included solar wind pressure and IMF |B| and Bz. Neither solar wind number density nor velocity added information so they were left out of the model. The maximum PACF was an acceptable 0.024 (Table 3, Model 6). This solar wind/IMF ARMAX model resulted in the somewhat lower validation correlation (0.611 vs. 0.675 of the source electron flux/SMEd model). In the timeplot of predictions, the solar wind/IMF ARMAX model tracks observations fairly well, although it does miss some of the higher peaks (Figure 6, pink line). The source electron flux/SMEd model predicts chorus more accurately.

In this model, again, *t-1* refers to a one day lag. Only |B|, Bz, and pressure (P) (units as given in Data section) were incorporated into the model as statistically significant inputs, along with an AR(1) and MA(1) term (based on chorus estimation from the model at a previous time step). Again, the automated procedure did not add lags previous to the one day lag in the numerator terms as previous lags did not improve the fit of the model. Unstandardized coefficients are given:

$$L4\ \widehat{Chorus}_t = -5.975 + \frac{.093\times|B|_t - .133\times|B|_{t-1}}{1-.682\times|B|_{t-2}} + \frac{-.250\times Bz_t}{1-.385\times Bz_{t-1}} + \frac{.288\times P_t}{1-.725\times P_{t-1}} + .965\times L4\ Chorus_{t-1} + .744\times\varepsilon_{t-1} \quad (8)$$

Using standardized regression coefficients, we compare the influence of the lagged effects over 14 days (Figure 8a). A more negative Bz (blue) is associated with higher chorus power. Increased |B| (orange) and pressure (green) also act to increase chorus power. All effects fall off rapidly over the course of a week.

4.4. Hourly ARMAX transfer function models

We created an hourly transfer function model of dayside chorus power at L4 with solar wind and IMF parameters as inputs. The hourly time series of chorus power is a more complicated function and was better fit by an ARMA (1,0,1)(1,0,1) model with AR and MA terms incorporating both a 1 hour and 24 hours lag of model-estimated chorus power and 1 hour and 24 hour moving average. Again, there was no need for differencing as there was no overall trend in the data. The highest magnitude term in the PACF of the ARMA hourly model without inputs was a statistically nonsignificant −0.046 at 3 hours (Model 7 of Table 3), thus this model removed the autocorrelation.

We entered all solar wind and IMF parameters and allowed the Expert Modeler function of SPSS to determine a best model. Solar wind number density was dropped from the final model as its presence did not increase the explanatory power. All terms remaining in the model are statistically significant. Note that the |B|, Bz, and pressure (P) terms act at a delay of 2 hours (i.e., they have the most influence on chorus power 2 hours after their measurement). Unstandardized coefficients are given in the equation, allowing calculation of predicted values from the observed transfer function predictor inputs:

$$L4\ \widehat{Chorus}_t = -3.894 + \frac{.026\times|B|_{t-2}}{1-.764\times|B|_{t-3}} + \frac{-.099\times Bz_{t-2}}{1-.654\times Bz_{t-3}} + \frac{.002\times V_t}{1-.567\times V_{t-1}} + \frac{.038\times P_{t-2}}{1-.521\times P_{t-3}} + .635\times L4\ Chorus_{t-1} + .056\times\varepsilon_{t-1} + .994\times L4\ Chorus_{t-24} + .980\times\varepsilon_{t-24} \quad (9)$$

While the model tracks the broad trends in the data, it is unable to predict the extremes highs and lows in chorus power seen by the DEMETER satellite. However, its validation correlation of 0.502 is an improvement over the 0.452 validation correlation from the hourly regression model with only lagged, observed chorus as an autoregressive input (Model 4).

Using standardized regression coefficients, we compare the influence of the lagged effects over 12 hours (Figure 8b). The effect of solar wind velocity (in yellow) begins in the first hour, while that of Bz (blue), |B| (orange), and pressure (green) all act beginning at a lag of 2 hours. All effects fall off rapidly. A more negative Bz is associated with higher chorus levels.

5. Discussion

5.1. Comparison of models

Serial autocorrelation in the response variable (chorus power) can be accounted for by the addition of ARMAX terms to the regression models. This correction for autocorrelation 1) gives us confidence that the effects of input parameters are being modeled correctly; 2) allows us to study the lagged effects of the predictors more accurately and efficiently; and 3) produces a predictive model that is not dependent on the availability of previous (lagged) values of chorus observations.

Our first models (Section 4a) use regression to predict chorus power levels using daily averages of source electron flux and SMEd measured with no lag and from the day before with no ARMA terms. Model 1 shows a significant influence of both inputs (flux and SMEd) at both selected lags. The prediction of chorus power in the test data set shows a reasonable validation correlation of 0.638. However, there is obvious serial autocorrelation. This may mean that the significant effects we see of flux and SMEd are not to be trusted. We can both reduce the serial autocorrelation and improve the validation correlation (to 0.766) by adding chorus power from the day before as an input to the model (Model 2). Although this is a reasonable predictive model, it depends on having chorus measurements available from the previous day. It is also not clear whether we have picked the best lags of the inputs to model chorus power, nor whether the correlation between lags of the same predictor may be causing difficulties in our model due to multicollinearity (a high correlation between predictors). Correlation between the same parameter at different lags may result in a model that is neither efficient nor true (Neter et al., 1985).

When we include all possible parameters in an hourly regression (including the lag of chorus power to correct for autocorrelation) this approach is even more awkward (Section 4b, Model 3). Testing only no lag, a one hour lag, and a 24 hour lag for each of the parameters gives a confusing model where some of the parameters show a significant influence at only some lags and some predictors show opposing effects with themselves at various lags. This is difficult to interpret and visualize, and the multicollinearity between the same parameter at different lags may still be a problem. Testing the influence of each parameter at greater lag times would be even more difficult. A further issue is that adding chorus observations from the previous hour did not correct the serial autocorrelation.

The ARMAX transfer function models (Section 4c and 4d) correct these problems. Serial autocorrelation of the residual errors is reduced to nonsignificant levels by the introduction of autoregressive and moving average terms that describe the time series behavior of chorus power measured at DEMETER. Once the model is trained, further inputs of chorus observations are not needed. In the source electron flux/SMEd transfer function model, both predictors are significant influences on chorus power on the same day, with more effect on the day before, and decreasing effect over days previous to that. In the daily solar wind/IMF parameter model, $|B|$, Bz, and pressure are the only parameters selected by the minimizing of the BIC. These variables are associated with increased chorus power beginning on the same day, with their effects decreasing over time. Both these models result in a similar validation correlation although the source electron flux/SMEd ARMAX model is somewhat more effective.

Although using known chorus observations from the previous time step fixes the autocorrelation problem in the daily model, using only the past ARMA behavior of chorus power (rather than regressing on the time lagged chorus values) resulted in predictions that were nearly as good as a model

dependent on lagged chorus observations. The results of this suggest that decent predictive models of other space weather parameters can be built that do not depend on obtaining continuous measurements of the variable to predict its own future behavior.

The hourly solar wind/IMF ARMAX model is not as successful at predicting chorus as the daily models. Predicting the erratic hourly chorus power measurements from DEMETER is difficult. Some smoothing of the data is necessary to obtain a better predictive model (Hyndman and Athanasopoulos, 2018), as was accomplished by using the daily averages, above. However, it is an improvement over the hourly regression model with all terms at several lags. The autocorrelation was successfully reduced in the ARMAX model, but not in the regression model, meaning we have confidence that the chosen parameters (solar wind velocity and pressure, IMF |B| and Bz) actually influence chorus power. The ARMAX hourly model also resulted in a higher validation correlation than the regression model including all parameters.

5.2. Drivers of chorus

Chorus wave production may fall into three steps. First, unstable distributions of source electrons are thought to drive chorus waves directly (Li et al., 2009; Meredith et al., 2001; 2003; Rodger et al., 2016; Su et al., 2014; Tsurutani & Smith, 1974). Second, these electrons are thought to be injected from the plasma sheet during substorm activity. (We measure substorm activity with the SMEd (darkside) auroral electrojet index which samples the magnetotail plasma sheet (Newell and Gjerloev, 2011).) Third, solar wind and IMF parameters drive substorm activity (Lyons et al., 2005; McPherron et al., 2009).

Single factor correlations (Figure 2) between chorus power and source electron flux or between chorus power and substorm activity (SMEd) suggest a linear relationship between both these factors and chorus power observed above L2. These factors do not have as much linear association with chorus power at L2, presumably because they do not penetrate to lower L shells. However, for SMEd, but not source electron flux, a curvilinear fourth-order polynomial fit gives a markedly higher correlation at L2. This suggests that substorm influences may penetrate more effectively to the lower L-shells during periods of high substorm activity.

The association is stronger between chorus power and source electron flux than with SMEd in these single factor correlations. This would be expected if source electrons were the direct driver of chorus waves and substorms were only an indirect driver through the generation of these source electrons. However, single factor simple correlations are not reliable in this situation for several reasons: 1) if the two predictors are themselves correlated, the relative influences may not be well represented by individual analyses; 2) serial autocorrelation of the response variable (chorus power) will result in inflation of the significance tests and the possibility of wrongly concluding that an explanatory variable is influential when it is, in fact, not; and 3) a serially autocorrelated response may be influenced by exogenous variables over many time steps which cannot be studied efficiently with correlation analysis. For these reasons, we test the relative influences of source electron flux and SMEd simultaneously using an ARMAX process.

In the ARMAX transfer function model, the influence of source electron flux and SMEd are equal on the first day, but the SMEd influence is higher on previous days, maintaining an influence up to two weeks

prior (Model 5; Figure 7). The source electron flux influence on chorus power observations is more immediate and a more fleeting phenomenon, likely because unstable distributions of source electrons act quickly as the direct driver of chorus. The SMEd influence is mostly indirect via the injection of these source electrons. The SMEd influence appears greater but this may only be because substorm activity continues to pump source electrons from the plasma sheet over a longer period of time while the action of the measured source electrons is quickly finished. It is also possible that substorms drive an additional, unaccounted for, process that drives the production of chorus waves, or that substorms are more reliably measured by the SMEd index than L4 source electrons are represented by the daily average flux observations made at the geosynchronous orbit of the LANL satellite ($L \approx 6.6$).

For both factors, the higher association with chorus power on the same day over the previous day in the simple correlations disappears in the transfer function model. The transfer function model, accounting for both serial autocorrelation of chorus power and the lagged effects of the inputs over time, gives us a clearer picture than simple correlation analysis.

Serial autocorrelation in a response variable is often seen in time series data. It may be an indication of physical persistence of the parameter or it may be the result of autocorrelation in either predictor or unmeasured variables that is not otherwise accounted for by the model. The autocorrelation in this system, therefore, could be an indication that chorus wave power does not dissipate quickly but remains detectable up to 24 hours later. Additionally, the chorus autocorrelation may be due to unmeasured processes that are themselves autocorrelated.

While this empirical model was developed to provide information on whether source electron flux or substorms are important in driving chorus, it could also be used as a predictive model. However, as source electron flux data is not always available as an input, we also built predictive ARMAX models using only the more readily available IMF and solar wind parameters: |B|, Bz, V, N, and P. The automatic selection technique used, which evaluates the BIC of all possible models and chooses the most parsimonious with the most explanatory power, gave us two models: a daily transfer function model with |B|, Bz, and pressure as inputs (Model 6), and an hourly model with |B|, Bz, pressure, and velocity as inputs (Model 7). The validation correlation of the daily solar wind/IMF model was lower than that of the daily source electron flux/SMEd model. Predictions from the daily solar wind/IMF model did not predict the highest chorus levels as well. The hourly ARMAX model (Model 7) was even less well correlated with observations. This may be because much of the variation seen in the hourly observations is due to fluctuations in the region the satellite is passing through rather than actual changes in chorus levels over time. As DEMETER is in a polar orbit, passing quickly through L shells in both the northern and southern hemispheres, it may be measuring large fluctuations in chorus power between hemispheres in any given hour. The hourly model was unable to forecast this variability with as much accuracy as the daily model.

Although inputs from IMF and solar wind sources may ultimately drive chorus activity via substorms and source electron flux generation, these inputs provide a weaker predictive model, even when chorus power levels are modelled as an ARMA process and even when all parameters are daily averaged.

6. Conclusions

1. Daily lower band chorus power is correlated with both radiation belt source electron flux (measured at L≈6) and substorm activity (SMEd index). Although the simple correlations are reasonably high at L3-L5 (up to 0.785 for source electron flux and 0.481 for SMEd), they are much lower at L2. Neither source electron flux nor substorm influences are associated with chorus power below L3.

2. A daily multiple regression model using both source electron flux and SMEd to predict chorus resulted in a validation correlation between observed and predicted values of 0.638. However, a regression or correlation will not accurately represent relationships if serial autocorrelation in the response variable (chorus power) is high. The autocorrelation of this regression model was unacceptably high at 0.310. A full hourly regression model with all variables entered had a validation correlation of 0.452, but the autocorrelation was too high at 0.444.

3. As one method to account for the problem of autocorrelation, we add an AR(1) term (average chorus power from the day before) to the regression. This lowers the autocorrelation to 0.103 and increases the validation correlation to 0.766. However, this model depends on having chorus observations from the day before to predict its value on the current day. Chorus measurements may not always be available.

4. We fit two daily autoregressive-moving average transfer function (ARMAX) models. These models used past behavior of chorus power (in the test dataset) and further inputs (source electron flux and SMEd, or solar wind and IMF parameters) to predict chorus power. Serial autocorrelation is greatly reduced, and validation correlations of 0.675 (for the source electron flux/SMEd ARMAX model) and 0.611 (for the solar wind/IMF model) show that there is good predictive ability.

5. The daily ARMAX models would be useful for predicting chorus power when chorus measurements are not available as predictions are based on the behavior of chorus power in the training set rather than on its value the day before.

6. While the source electron flux/SMEd ARMAX model tracked chorus power changes well in a timeplot, the solar wind/IMF ARMAX predictions were somewhat less accurate. If source electon flux and SMEd are not available as inputs, the solar wind and IMF model could be used instead, but with a loss of accuracy.

7. Hourly models, either regression or ARMAX, were not able to predict chorus as effectively.

8. ARMAX models are better able to analyze the time lagged effects of predictors than regression. From the ARMAX model, we were able to determine that the action of source electron flux on chorus power levels is highest in the first few days. Substorm activity

(measured by SMEd) has more long-lasting effects, presumably due to its continued pumping of source electrons from the plasma sheet.

9. Not all solar wind and IMF parameters are useful in predicting chorus power. The lowest BIC was obtained for ARMAX models containing IMF |B| and Bz, and solar wind pressure (in the daily model) and |B|, Bz, pressure, and solar wind velocity (in the hourly model). Number density (N) was never chosen by the automated selection technique.

7. Acknowledgements

We thank two anonymous reviewers for improving this manuscript. We also thank R. Gamble for preparing and J.-J. Berthelier for providing DEMETER ICE data. All DEMETER are now available at the CDPP (Centre de données de la Physique des Plasmas) website: cdpp-archive.cnes.fr. Source electron flux data were obtained from Los Alamos National Laboratory (LANL) geosynchronous energetic particle instruments (contact: G. D. Reeves). The SMEd index is available from SuperMAG (http://supermag.jhuapl.edu/ , Principal Investigator Jesper Gjerloev), derived from magnetometer data from Intermagnet; USGS, Jeffrey J. Love; CARISMA, PI Ian Mann; CANMOS; The S-RAMP Database, PI K. Yumoto and Dr. K. Shiokawa; The SPIDR database; AARI, PI Oleg Troshichev; The MACCS program, PI M. Engebretson, Geomagnetism Unit of the Geological Survey of Canada; GIMA; MEASURE, UCLA IGPP and Florida Institute of Technology; SAMBA, PI Eftyhia Zesta; 210 Chain, PI K. Yumoto; SAMNET, PI Farideh Honary; The institutes who maintain the IMAGE magnetometer array, PI Eija Tanskanen; PENGUIN; AUTUMN, PI Martin Connors; DTU Space, PI Dr. Rico Behlke; South Pole and McMurdo Magnetometer, PI's Louis J. Lanzerotti and Alan T. Weatherwax; ICESTAR; RAPIDMAG; PENGUIn; British Artarctic Survey; McMAC, PI Dr. Peter Chi; BGS, PI Dr. Susan Macmillan; Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (IZMIRAN); GFZ, PI Dr. Juergen Matzka; MFGI, PI B. Heilig; IGFPAS, PI J. Reda; University of L'Aquila, PI M. Vellante. IMF Bz, Dst, and solar wind V, N, and P data are available from Goddard Space Flight Center Space Physics Data Facility at the OMNIWeb data website (http://omniweb.gsfc.nasa.gov/html/ow_data.html). Work at Augsburg University was supported by NSF grant AGS-1651263.

8. Literature Cited

Balikhin, M. A., R. J. Boynton, S. N. Walker, J. E. Borovsky, S. A. Billings, and H. L. Wei (2011), Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit, Geophysical Research Letters, 38, L18105, doi:10.1029/2011GL048980

662 Berthelier, J.J., M. Godefroy, F. Leblanc, M. Malingre, M. Menvielle, D. Lagoutte, J.Y. Brochot, F. Colin, F.
663 Elie, C. Legendre, P. Zamora, D. Benoist, Y. Chapuis, J. Artru, R. Pfaff (2006), ICE, the electric field
664 experiment on DEMETER, Planetary and Space Science 54, 456–471, doi:10.1016/j.pss.2005.10.016
665
666 Borovsky, J. E., and M. H. Denton (2014), Exploring the cross correlations and autocorrelations of the
667 ULF indices and incorporating the ULF indices into the systems science of the solar wind-driven
668 magnetosphere, Journal of Geophysical Research Space Physics, 119, doi:10.1002/2014JA019876
669
670 Bortnik, J., R. M. Thorne, T. P. O'Brien, J. C. Green, R. J. Strangeway, Y. Y. Shprits, and D. N. Baker (2006),
671 Observation of two distinct, rapid loss mechanisms during the 20 November 2003 radiation belt dropout
672 event, Journal of Geophysical Research, 111, A12216, doi:10.1029/2006JA011802
673
674 Bortnik, J. and R.M. Thorne (2007), The dual role of ELF/VLF chorus waves in the acceleration and
675 precipitation of radiation belt electrons, Journal of Atmospheric and Solar-Terrestrial Physics, 69 (2007)
676 378–386, doi:10.1016/j.jastp.2006.05.030.
677

678 Boynton, R. J., M. A. Balikhin, S. A. Billings, and O. A. Amariutei (2013a), Application of nonlinear
679 autoregressive moving average exogenous input models to Geospace: Advances in understanding and
680 space weather forecasts, Annales Geophysicae, 31, 1579–1589.

681 Boynton, R. J., M. A. Balikhin, S. A. Billings, G. D. Reeves, N. Ganushkina, M. Gedalin, O. A. Amariutei, J. E.
682 Borovsky, and S. N. Walker (2013b), The analysis of electron fluxes at geosynchronous orbit employing a
683 NARMAX approach, Journal of Geophysical Research, Space Physics, 118, 1500–1513,
684 doi:10.1002/jgra.50192.

685 Gjerloev, J. W. (2012), The SuperMAG data processing technique, Journal of Geophysical Research, 117,
686 A09213, doi:10.1029/2012JA017683

687 Hendry, A. T., C. J. Rodger, and M. A. Clilverd (2017), Evidence of sub-MeV EMIC-driven electron
688 precipitation, Geophysical Research Letters, 44, 1210–1218, doi:10.1002/2016GL071807

689 Hikishima, M., Y. Omura, and D. Summers (2010), Microburst precipitation of energetic electrons
690 associated with chorus wave generation, Geophysical Research Letters, 37, L07103,
691 doi:10.1029/2010GL042678

692 Horne, R. B. and R. M. Thorne (1998), Potential waves for relativistic electron scattering and stochastic
693 acceleration during magnetic storms, Geophysical Research Letters, 25, 3011–3014, doi:
694 10.1029/98gl01002

695 Horne, R. B., and R. M. Thorne (2003), Relativistic electron acceleration and precipitation during
696 resonant interactions with whistler-mode chorus, Geophysical Research Letters, 30, 1527,
697 doi:10.1029/2003GL016973, 10

698 Hyndman, R.J. and G. Athanasopoulos (2018) Forecasting: Principles and Practice, 2nd ed., OTexts,
699 Heathmont, Victoria, Australia 291 pp.

700 Lam, M. M. R. B. Horne, N. P. Meredith, S. A. Glauert, T. Moffat-Griffin, and J. C. Green (2010), Origin of
701 energetic electron precipitation >30 keV into the atmosphere, Journal of Geophysical Research, 115,
702 A00F08, doi:10.1029/2009JA014619

Li, L., J. B. Cao, and G. C. Zhou (2005), Combined acceleration of electrons by whistler-mode and compressional ULF Pc5 turbulences near the geosynchronous orbit, Journal of Geophysical Research, 110, A03203, doi:10.1029/2004JA010628

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, Geophysical Research Letters, 36, L09104, doi:10.1029/2009GL037595

Lorentzen, K. R., J. B. Blake, U. S. Inan, and J. Bortnik (2001), Observations of relativistic electron microbursts in association with VLF chorus, Journal of Geophysical Research, 106, 6017.

Lyons, L. R., D.-Y. Lee, R. M. Thorne, R. B. Horne, and A. J. Smith (2005), Solar wind-magnetosphere coupling leading to relativistic electron energization during high-speed streams, Journal of Geophysical Research,110, A11202, doi:10.1029/2005JA011254

Makridakis, S.G., S. C. Wheelwright, and R. J. Hyndman (1998) Forecasting: Methods and Applications, 3rd ed., John Wiley and Sons, New York, NY, 652 pp.

McPherron, R.L., D.N.Baker, N.U.Crooker (2009), Role of the Russell–McPherron effect in the acceleration of relativistic electrons, Journal of Atmospheric and Solar-Terrestrial Physics, 71(2009)1032–1044.

Meredith, N.P., R.B. Horne, and R. R. Anderson (2001), Substorm dependence of chorus amplitudes: Implications for the acceleration of electrons to relativistic energies, Journal of Geophysical Research, 106,13,165-13,178, doi: 10.1029/2000JA900156

Meredith, N. P., R. B. Horne, R. H. A. Iles, R. M. Thorne, D. Heynderickx, and R. R. Anderson (2002), Outer zone relativistic electron acceleration associated with substorm-enhanced whistler mode chorus, Journal of Geophysical Research, 107(A7), 1144, 10.1029/2001JA900146

Meredith, N. P., R. B. Horne, R. M. Thorne, and R. R. Anderson (2003), Favored regions for chorus-driven electron acceleration to relativistic energies in the Earth's outer radiation belt, Geophysical Research Letters, 30(16), 1871, doi:10.1029/2003GL017698

Millan, R.M. and R.M. Thorne (2007), Review of radiation belt relativistic electron losses, Journal of Atmospheric and Solar-Terrestrial Physics, 69 (2007) 362–377, doi:10.1016/j.jastp.2006.06.019

Neter, J., W. Wasserman, and M. H. Kutner (1985), Applied Linear Statistical Models, Richard D. Irwin, Inc., Homewood, Ill.

Neal, J. J., C. J. Rodger, M. A. Clilverd, N. R. Thomson, T. Raita, and Th. Ulich (2015), Long-term determination of energetic electron precipitation into the atmosphere from AARDDVARK subionospheric VLF observations, J. Geophys. Res., 120, 2194–2211, doi:10.1002/2014JA020689

Newell, P. T., and J. W. Gjerloev (2011), Evaluation of SuperMAG auroral electrojet indices as indicators of substorms and auroral power, J. Geophys. Res., 116, A12211, doi:10.1029/2011JA016779.

Orlova, K. G. and Y.Y. Shprits (2010), Dependence of pitch-angle scattering rates and loss timescales on the magnetic field model, Geophysical Research Letters, 37, L05105, doi:10.1029/2009GL041639

739 Pankratz, A. (1991) Forecasting with Dynamic Regression Models, John Wiley and Sons, New York, NY,
740 386 pp.

741 Reeves, G. D., S. K. Morley, R. H. W. Friedel, M. G. Henderson, T. E. Cayton, G. Cunningham, J. B. Blake, R.
742 A. Christensen, and D. Thomsen (2011), On the relationship between relativistic electron flux and solar
743 wind velocity: Paulikas and Blake revisited, J. Geophys. Res., 116, A02213, doi:10.1029/2010JA015735

744 Rodger, C.J., K.Cresswell-Moorcock, and M.A.Clilverd (2016), Nature's Grand Experiment: Linkage
745 between magnetospheric convection and the radiation belts, Journal of Geophysical Research, Space
746 Physics, 121,171–189, doi:10.1002/2015JA02153

747 Santolik, O., J. Chum, M. Parrot, D. A. Gurnett, J. S. Pickett, and N. Cornilleau-Wehrlin (2006),
748 Propagation of whistler mode chorus to low altitudes: Spacecraft observations of structured ELF hiss, J.
749 Geophys. Res., 111, A10208, doi:10.1029/2005JA011462

750 Schwarz, G. E. (1978), Estimating the dimension of a model, Annals of Statistics, 6 (2): 461–464,
751 doi:10.1214/aos/1176344136

752 Simms, L., M. Engebretson, M. Clilverd, C. Rodger, M. Lessard, J. Gjerloev, and G. Reeves (2018a), A
753 distributed lag autoregressive model of geostationary relativistic electron fluxes: Comparing the
754 influences of waves, seed and source electrons, and solar wind Inputs, Journal of Geophysical Research:
755 Space Physics, 123(5), 3646–3671, doi:10.1029/2017ja025002

756 Simms, L.E., M. J. Engebretson, M. A. Clilverd, C. J. Rodger, M. R. Lessard, and G. D. Reeves (2018b),
757 Nonlinear and synergistic effects of ULF Pc5, VLF chorus, and EMIC waves on relativistic electron flux at
758 geosynchronous orbit. Journal of Geophysical Research: Space Physics doi:10.1029/2017JA025003

759 Su, Z., et al. (2014), Intense duskside lower band chorus waves observed by Van Allen Probes:
760 Generation and potential acceleration effect on radiation belt electrons, Journal of Geophysical
761 Research, Space Physics, 119, 4266–4273, doi:10.1002/2014JA019919

762 Summers, D., R. M. Thorne, and F. Xiao (1998), Relativistic theory of wave-particle resonant diffusion
763 with application to electron acceleration in the magnetosphere, Journal of Geophysical Research, 103,
764 20,487.

765 Summers, D., B. Ni, and N. P. Meredith (2007), Timescales for radiation belt electron acceleration and
766 loss due to resonant wave-particle interactions: 2. Evaluation for VLF chorus, ELF hiss, and
767 electromagnetic ion cyclotron waves, Journal of Geophysical Research, 112, A04207,
768 doi:10.1029/2006JA011993

769 Thorne, R. M., W. Li, B. Ni, Q. Ma, J. Bortnik, L. Chen, D. N. Baker, H. E. Spence, G. D. Reeves, M. G.
770 Henderson, C. A. Kletzing, W. S. Kurth, G. B. Hospodarsky, J. B. Blake, J. F. Fennell, S. G. Claudepierre, and
771 S. G. Kanekal (2013), Rapid local acceleration of relativistic radiation-belt electrons by magnetospheric
772 chorus, Nature, 504, 411, doi:10.1038/nature12889

773 Tsurutani, B. T., and E. J. Smith (1974), Postmidnight chorus: A substorm phenomenon, Journal of
774 Geophysical Research, 79, 118–127, doi:10.1029/JA079i001p00118

775 Xiao, F., C. Yang, Z. Su, Q. Zhou, Z. He, Y. He, D.N. Baker, H.E. Spence, H.O. Funsten, and J.B. Blake (2015),
776 Wave-driven butterfly distribution of Van Allen belt relativistic electrons, Nature Communications, doi:
777 10.1038/ncomms9590

778 Zhima, Z., J. Cao, W. Liu, H. Fu, J. Yang, X. Zhang, and X. Shen (2013), DEMETER observations of high-
779 latitude chorus waves penetrating the plasmasphere during a geomagnetic storm, Geophys. Res. Lett.,
780 40, 5827–5832, doi:10.1002/2013GL058089

781 Table 1. Comparison of regression models using source electron flux and SMEd (Lag0 hours and Lag24 hours) and/or solar wind and IMF
782 parameters (Lag0, Lag1, and Lag24) as input parameters. Models 1 and 4 are uncorrected for serial autocorrelation. Models 2, 3, and 5 attempt
783 to correct for autocorrelation by adding observed chorus at the previous time step as an input parameter.

784 R^2, the prediction efficiency, is the fraction of variation in the training set data explained by the model. The validation correlation is the
785 correlation between observed chorus and that predicted by the model. The maximum PACF is the highest autocorrelation in the residual errors
786 from the model.

Model	Model type	Time Step	Lag chorus	Source electron flux and SMEd	Solar wind and IMF parameters	R^2 Variation explained by model	Validation correlation	Maximum PACF	Figure	Equation
1	Regression	Daily		✓		0.644	0.638	0.314 (day 1)	3ab	5
2	Regression	Daily	✓	✓		0.710	0.766	0.104 (day 4)	3cd, 6	6
3	Regression	Daily	✓	✓	✓	0.775	0.823	0.082 (day 4)	--	--
4	Regression	Hourly	✓	✓	✓	0.562	0.452	0.453 (hour 1)	4	Table 2

787
788

789 Table 2. Unstandardized coefficients of the hourly regression model (Model 4 of Table 1) predicting chorus power from predictors at the current
790 time (Lag0), from 1 h previous (Lag1), and from 24 h previous (Lag24). Significant predictors (p-value < 0.05) are denoted with an asterisk.

	Lag0	Lag1	Lag24
Constant	-3.156*		
|B|	-0.005	0.021*	0.050
Bz	0.030*	-0.080*	-0.005
Source Flux	0.258*	--	0.185*
SMEd	0.0002	--	-0.0004*
N	0.017*	-0.012*	-0.003
V	0.002*	-0.0005	-0.0004*
P	0.030*	0.004	0.010
Chorus power	--	0.580*	--

791

Table 3. Comparison of ARMAX models. These models correct for serial autocorrelation by adding AR and MA terms that model the time series behavior of chorus rather than inputting actual chorus observations. Daily models include an AR(1) term and an MA term (either 1 or 4). The hourly model contains an AR(1) and MA(1) term as well as a "seasonal" (24 hour) AR(1) and MA(1) term. Source electron flux and SMEd and/or solar wind and IMF parameters are included as transfer function inputs.

R^2, the prediction efficiency, is the fraction of variation in the training set data explained by the model. Validation correlation is the correlation between observed chorus and that predicted by the model. The maximum PACF is the highest autocorrelation in the residual errors from the model.

Model	Model type (p,d,q)	Time Step	Source electron flux and SMEd	Solar wind and IMF parameters	R^2 Variation explained by model	Validation correlation	Maximum PACF	Figure	Equation
5	ARMAX(1,0,4)	Daily	✓		0.732	0.675	0.081 (day 4)	5, 7	7
6	ARMAX(1,0,1)	Daily		✓	0.734	0.611	0.024 (day 1)	6, 8a	8
7	ARMAX (1,0,1)(1,0,1)	Hourly		✓	0.628	0.502	-0.046 (hour 3)	8b	9

Figure 1. Timeplots of DEMETER L4 lower band chorus power (gray), source electron flux (blue), and SMEd (red) over Sept. 2006 at L4 (L=4-4.99). Observations are standardized to produce z-scores by subtracting the overall mean and dividing by the standard deviation of each variable.

Figure 2. Correlation of lower band chorus DEMETER satellite observations (0.1-0.5 fce) with source electron flux or SMEd at L 2-5. a. source electron flux correlated with dayside DEMETER observations, b. source electron flux correlated with nightside DEMETER observations, c. SMEd with dayside DEMETER observations, and d. SMEd with nightside DEMETER observations. Chorus observations on a given day are correlated with source electron flux or SMEd daily average on the same day (red) and from 24 hours before (black). Simple linear correlations are shown in black or red, correlations with a fourth-order polynomial model in gray or pink.

Figure 3. Daily averaged lower band chorus power ($\log10(\mu V^2/m^2/Hz)$) predicted by source electron flux and SMEd (top row; Model 1 of Table 1), or by source electron flux, SMEd, and chorus on the day before (bottom row; Model 2 of Table 1). a. Scatterplot of observed L4 lower band chorus power vs. that predicted by regression model using source electron flux and SMEd as predictors. Validation correlation is 0.638. b. Partial autocorrelation function (PACF) of the residual errors of the model. Autocorrelation at a one day lag of 0.310 is unacceptably high and indicates uncorrected serial autocorrelation. c. Scatterplot of observed chorus power and that predicted by regression model using source electron flux, SMEd, and chorus power on the previous day as predictors. Validation correlation is higher with previous chorus added as a predictor (r=0.766). d. The maximum partial autocorrelation of the residuals is 0.103 indicating that some of the serial autocorrelation has been removed by the addition of previous day's chorus.

Figure 4. Full regression model predicting hourly dayside L4 chorus power ($\log10(\mu V^2/m^2/Hz)$) using source electron flux, SMEd, IMF |B| and Bz, and solar wind density (N), velocity (V), and pressure (P) as well as chorus from the previous hour to control for serial autocorrelation (Model 4 of Table 1). a. Standardized coefficients from the multiple regression analysis. Predictors are measured at the same time as chorus (lag 0), an hour before the chorus measurement (lag 1), and 24 hours before (lag 24). Only lag 0 and lag 24 of source electron flux and SMEd are used as they are daily measurements. Statistically significant effects (p-value < 0.05) are shown in gray; nonsignificant effects are in white. Chorus at lag 1 is included as a predictor but not shown in this figure. Its standardized regression coefficient was much higher than the other predictors at 0.577. b. Partial autocorrelations of the residual errors of the regression model show that there is still significant autocorrelation of chorus at lag1 (0.453) despite the addition of lag 1 chorus as a predictor. c. Scatterplot of observed chorus and that predicted by the regression model (validation correlation = 0.452).

Figure 5. Predicting daily L4 chorus power ($\log10(\mu V^2/m^2/Hz)$) using ARMAX transfer function model with source electron flux and SMEd as inputs (Model 5 of Table 3). a. Validation correlation of observed chorus and that predicted by the ARMAX transfer function model (r = 0.675). b. Partial autocorrelations of the residual errors of this ARMAX model. The serial autocorrelation is reduced to an acceptable level.

Figure 6. Timeplot of chorus power ($\log10(\mu V^2/m^2/Hz)$) predicted by ARMAX model using source flux and SMEd as inputs (green; Model 5 of Table 3), an ARMAX full model using solar wind and IMF parameters

847 |B|, Bz, and pressure (pink; Model 6 of Table 3), and the regression model using source electron flux,
848 SMEd, and previous day's chorus (blue line; Model 4 of Table 1) over Feb-May of 2009.
849
850 Figure 7. Standardized cumulative effects on L4 lower band chorus power over 0-14 days previous of
851 source electron flux (black bars) and SMEd (gray bars) (Model 5 of Table 3).
852
853 Figure 8. Standardized cumulative effects on L4 lower band chorus power. a. over 0-14 days using the
854 daily ARMAX(1,0,1) model (Model 6) with IMF |B| (orange), IMF Bz (blue), and solar wind pressure
855 (green) as predictors, and b. over 0-12 hours using the hourly ARMAX(1,0,1)(1,0,1) model (Model 7)
856 using solar wind velocity (white), IMF |B| (orange), IMF Bz (blue), and pressure (green) as predictors.
857

Figure 1.

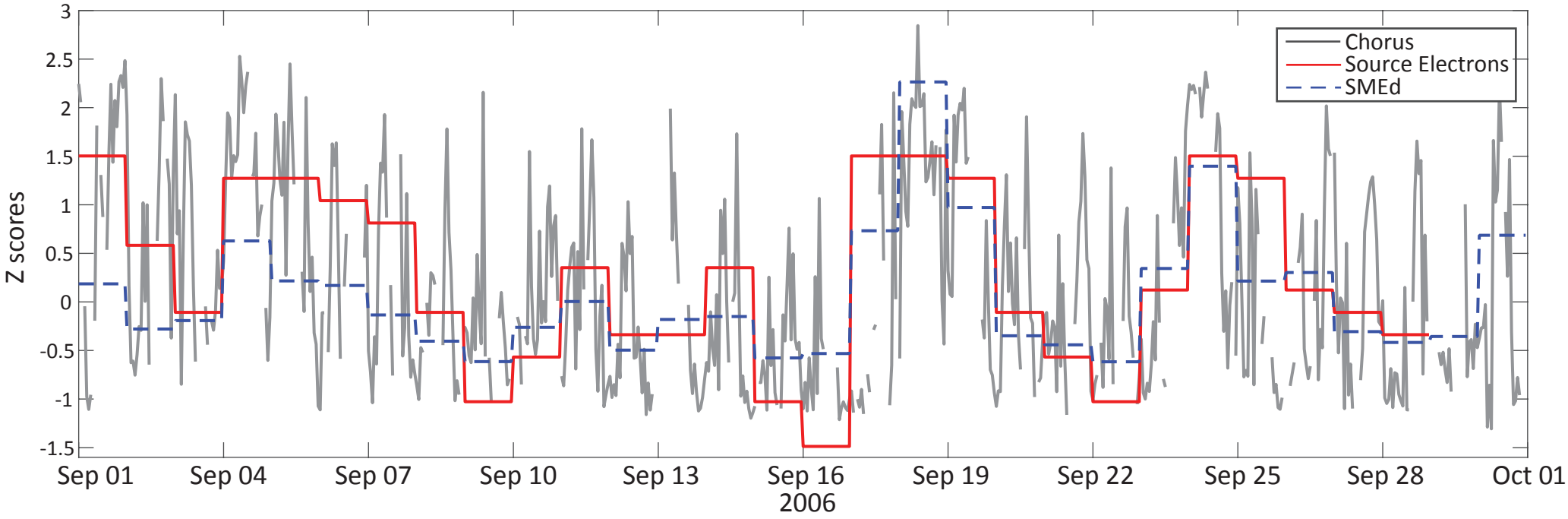

Chorus
Source Electrons
SMEd
Z scores
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
Sep 01
Sep 04
Sep 07
Sep 10
Sep 13
Sep 16
Sep 19
Sep 22
Sep 25
Sep 28
Oct 01
2006

Figure 2.

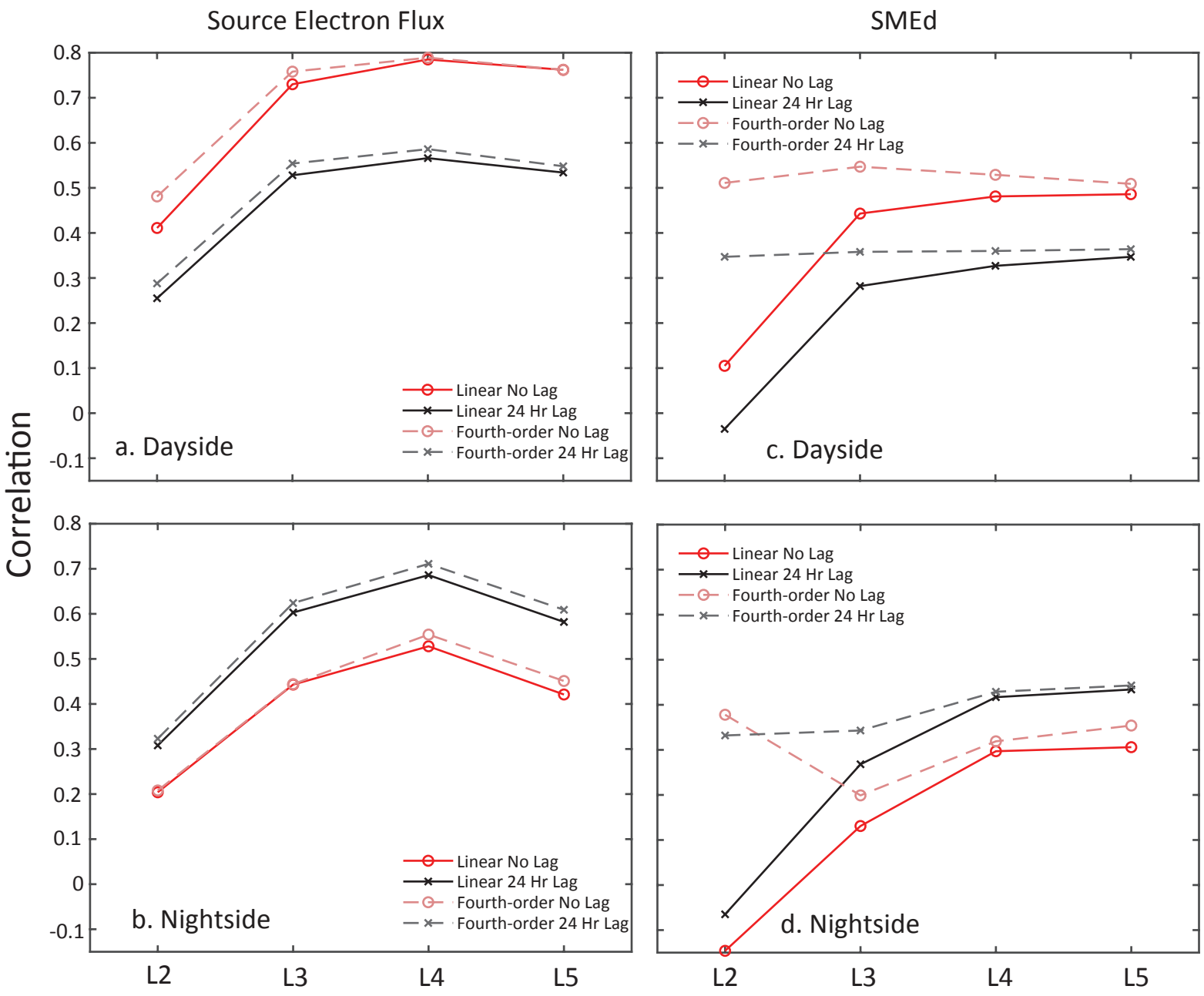
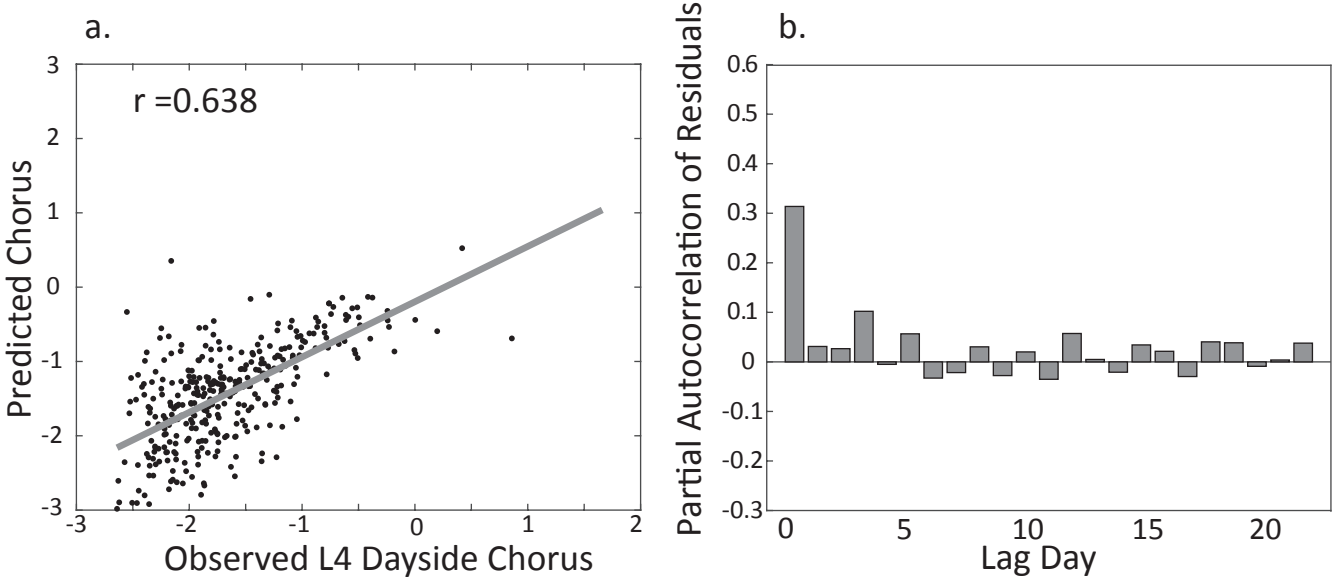

Source Electron Flux
SMEd
Correlation
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
a. Dayside
b. Nightside
c. Dayside
d. Nightside
Linear No Lag
Linear 24 Hr Lag
Fourth-order No Lag
Fourth-order 24 Hr Lag
L2
L3
L4
L5

Figure 3.

Chorus predicted by source electron flux and SMEd regression model



Chorus predicted by source electron flux, SMEd, and chorus from previous day

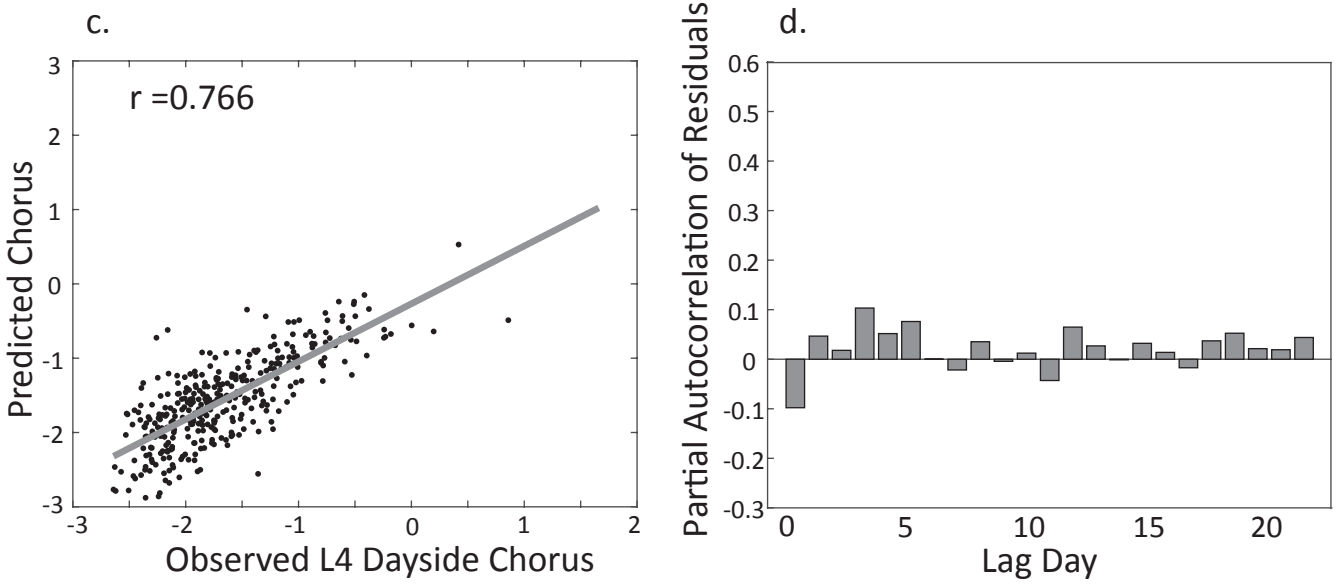
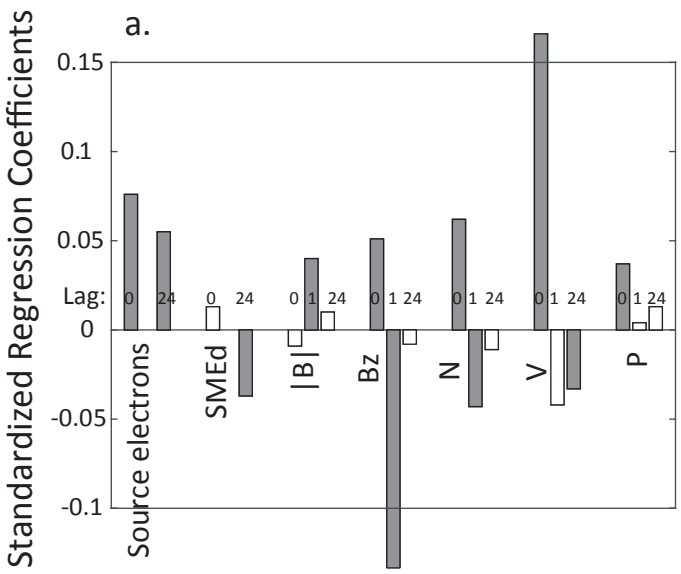
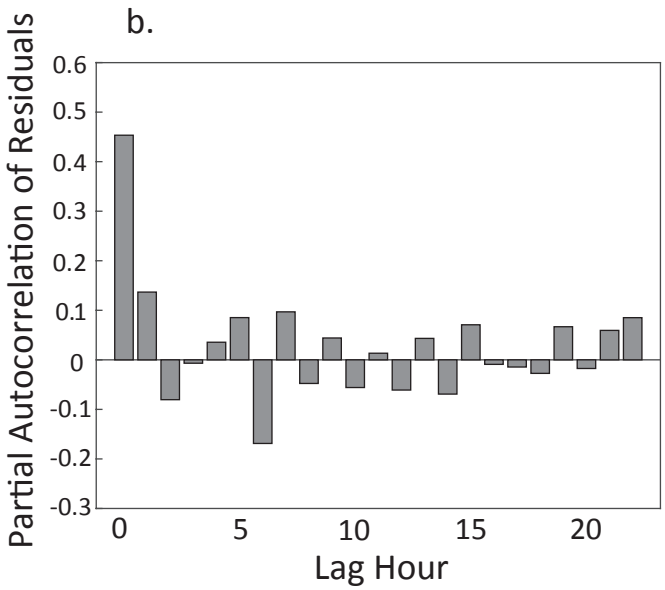


Figure 4.


a.
Standardized Regression Coefficients
0.15
0.1
0.05
0
-0.05
-0.1
Lag:
0 24
0 24
0 1 24
0 1 24
0 1 24
0 1 24
0 1 24
Source electrons
SMEd
|B|
Bz
N
V
P


b.
Partial Autocorrelation of Residuals
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0
5
10
15
20
Lag Hour

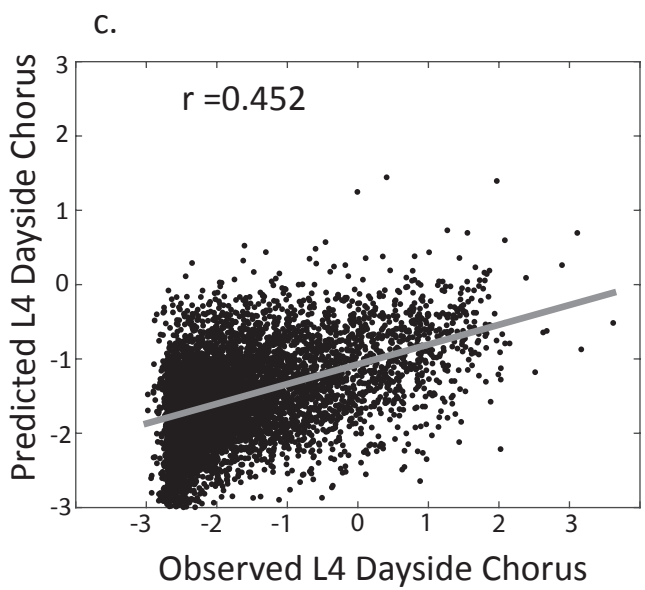

c.
r =0.452
Predicted L4 Dayside Chorus
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
Observed L4 Dayside Chorus

Figure 5.

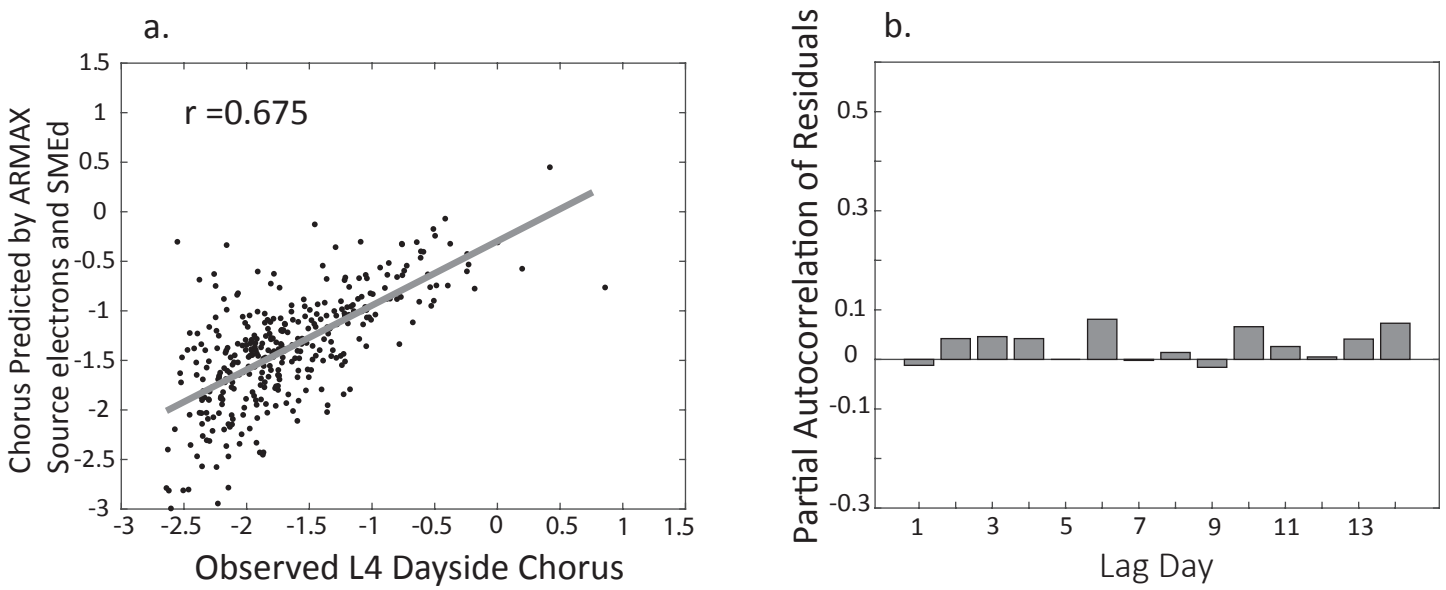

a.
r =0.675
Chorus Predicted by ARMAX Source electrons and SMEd
Observed L4 Dayside Chorus
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
b.
Partial Autocorrelation of Residuals
0.5
0.3
0.1
0
-0.1
-0.3
1
3
5
7
9
11
13
Lag Day

Figure 6.

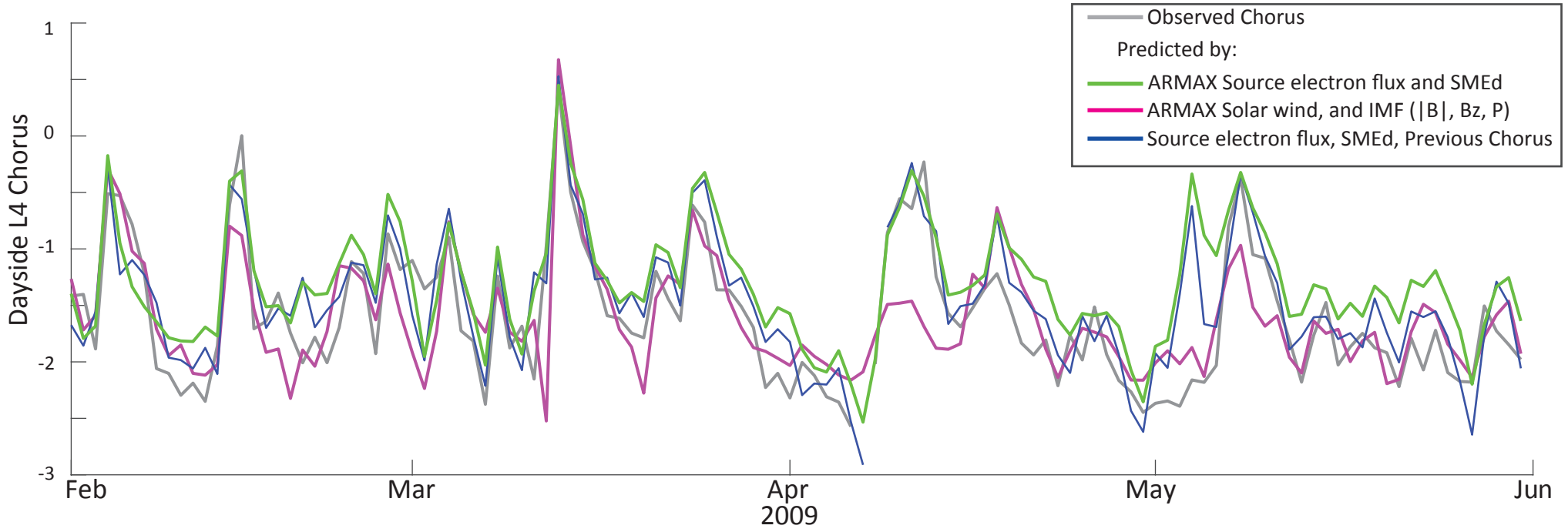

Observed Chorus
Predicted by:
ARMAX Source electron flux and SMEd
ARMAX Solar wind, and IMF (|B|, Bz, P)
Source electron flux, SMEd, Previous Chorus
Dayside L4 Chorus
1
0
-1
-2
-3
Feb
Mar
Apr
2009
May
Jun

Figure 7.

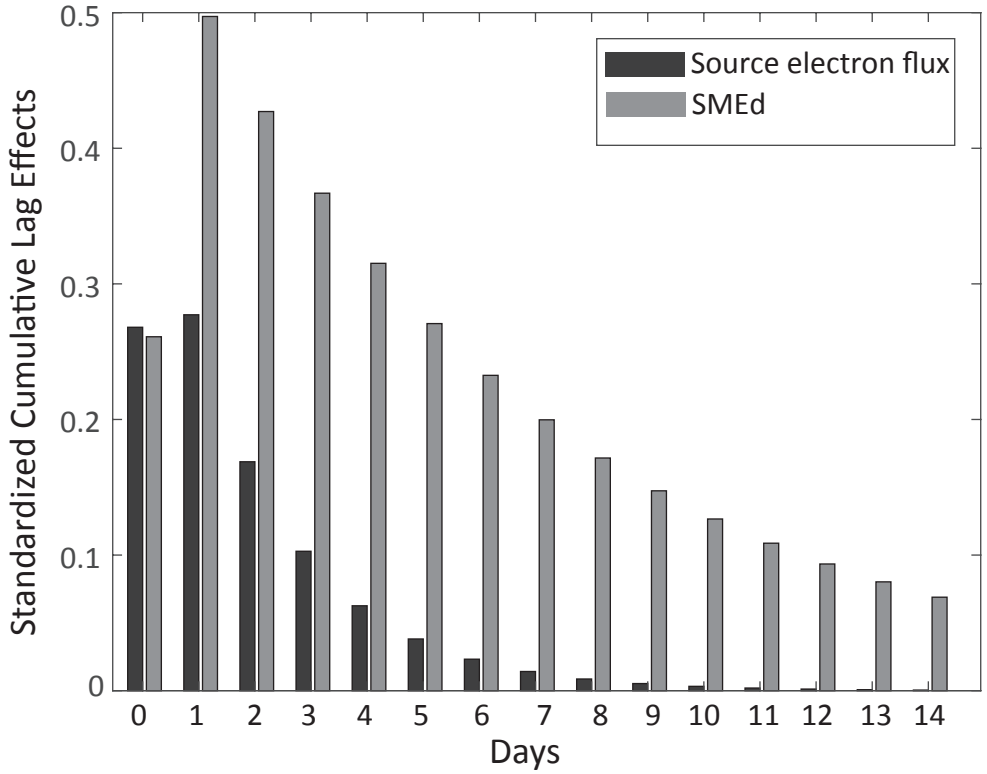
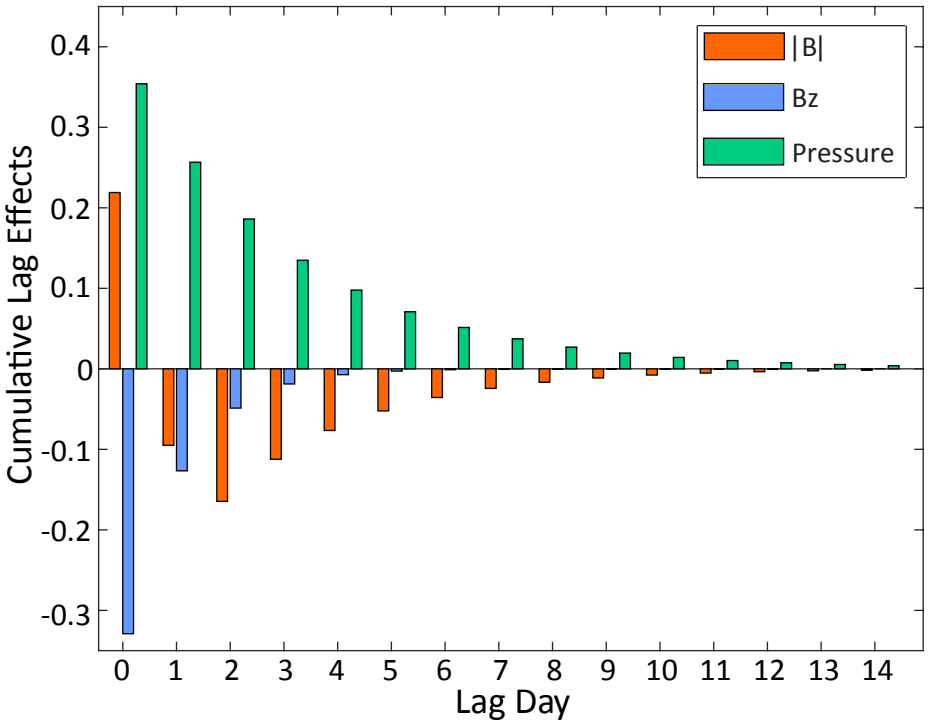

Source electron flux
SMEd
Standardized Cumulative Lag Effects
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Days

Figure 8.

a. Daily Model



b. Hourly model