AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Space Physics

RESEARCH ARTICLE
10.1002/2016JA023746

Key Points:
- Electron temperature near midnight geosynchronous orbit varies from 0.1 to 10 keV
- Electron temperature increases with increasing southward IMF B_z and solar wind speed
- Higher electron temperature when solar wind ULF powers are higher

Correspondence to:
C.-P. Wang,
cat@atmos.ucla.edu

Citation:
Wang, C.-P., H.-J. Kim, C. Yue, J. M. Weygand, T.-S. Hsu, and X. Chu (2017), Effects of solar wind ultralow-frequency fluctuations on plasma sheet electron temperature: Regression analysis with support vector machine, *J. Geophys. Res. Space Physics*, *122*, doi:10.1002/2016JA023746.

Received 29 NOV 2016
Accepted 20 MAR 2017
Accepted article online 23 MAR 2017

Effects of solar wind ultralow-frequency fluctuations on plasma sheet electron temperature: Regression analysis with support vector machine

Chih-Ping Wang[1], Hee-Jeong Kim[1], Chao Yue[1,2], James M. Weygand[3], Tung-Shin Hsu[3], and Xiangning Chu[3]

[1]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, California, USA, [2]University Corporation for Atmospheric Research, Boulder, Colorado, USA, [3]Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, California, USA

Abstract To investigate whether ultralow-frequency (ULF) fluctuations from 0.5 to 8.3 mHz in the solar wind and interplanetary magnetic field (IMF) can affect the plasma sheet electron temperature (T_e) near geosynchronous distances, we use a support vector regression machine technique to decouple the effects from different solar wind parameters and their ULF fluctuation power. T_e in this region varies from ~0.1 to 10 keV with a median of 1.3 keV. We find that when the solar wind ULF power is weak, T_e increases with increasing southward IMF B_z and solar wind speed, while it varies weakly with solar wind density. As the ULF power becomes stronger during weak IMF B_z (~0) or northward IMF, T_e becomes significantly enhanced, by a factor of up to 10. We also find that mesoscale disturbances in a time scale of a few to tens of minutes as indicated by *AE* during substorm expansion and recovery phases are more enhanced when the ULF power is stronger. The effect of ULF powers may be explained by stronger inward radial diffusion resulting from stronger mesoscale disturbances under higher ULF powers, which can bring high-energy plasma sheet electrons further toward geosynchronous distance. This effect of ULF powers is particularly important during weak southward IMF or northward IMF when convection electric drift is weak.

1. Introduction

The nightside plasma sheet is the main particle source for ring current and radiation belt. Variations in the plasma sheet electron temperature are thus important for development of relativistic electrons. Solar wind energy input into the magnetosphere, which depends on types of solar wind/interplanetary magnetic field (IMF) structure, is one of the major processes to accelerate plasma sheet particles. This external energy input controls the magnetospheric convection and thus large-scale electron inward motion and adiabatic energization due to $E \times B$ drift [*Wang et al.*, 2011]. Previous observational studies have identified several solar wind/IMF parameters as primary external drivers of the plasma sheet temperature. The plasma sheet temperature is correlated with the solar wind velocity [*Borovsky et al.*, 1998; *Tsyganenko and Mukai*, 2003; *Wang et al.*, 2007; *Dubyagin et al.*, 2016]. Plasma sheet temperature is lower for positive IMF B_z than negative IMF B_z but with the effects depending strongly on the history of IMF B_z [e.g., *Terasawa et al.*, 1997; *Øieroset et al.*, 2003; *Wing et al.*, 2005; *Wang et al.*, 2010]. However, most of these previous studies focused only on plasma sheet ion temperature and our understanding of the electron temperature remains limited.

It has been observed that solar wind energy input can also be affected by fluctuations of solar wind/IMF parameters in the ultralow-frequency (ULF, in a time scale of a few to tens of minutes) range [e.g., *Kim et al.*, 2009]. These ULF fluctuations are often associated with Alfvénic waves within the interplanetary medium [*Belcher and Davis*, 1971]. Stronger IMF fluctuations are found to elevate polar-cap convection strength consisting of transient but larger enhancements [*Kim et al.*, 2011, 2012], as well as to cause more frequent substorm occurrence [*Tsurutani et al.*, 1990; *Lyons et al.*, 2009]. Stronger ULF fluctuations in solar wind velocity and density can enhance ULF wave activity within the magnetosphere [*Kessel*, 2008]. Compared with coronal mass ejection-driven storms, solar wind ULF fluctuations during corotating interaction region-driven storms are larger and relativistic electron fluxes are higher [*Borovsky and Denton*, 2006; *Potapov*, 2013]. These observations thus suggest that solar wind ULF fluctuations may also affect the acceleration of plasma sheet electrons.

Besides the solar wind/IMF conditions, plasma sheet electrons are also affected by plasma sheet mesoscale disturbances in a time scale of a few to several tenths of minutes, such as bursty bulk flows, depolarization fronts, and substorms. Electrons can be accelerated within these mesoscale disturbances [e.g., *Liang et al.*, 2014; *Birn et al.*, 2014; *Grigorenko et al.*, 2016]. These disturbances could also lead to earthward penetration and energization of plasma sheet particles which are more localized than those by large-scale $E \times B$ drift [e.g., *Gabrielse et al.*, 2016; *Yang et al.*, 2016]. These mesoscale disturbances are stronger when *AE* is higher [e.g., *Angelopoulos et al.*, 1994; *Wang et al.*, 2009]. Also, as *AE* becomes higher, the plasma sheet electron temperature is found to be higher [e.g., *Wang et al.*, 2009] and the electron Alfvén layers to penetrate further earthward [e.g., *Korth et al.*, 1999; *Jiang et al.*, 2011]. Furthermore, it is well known that *AE* level is closely correlated with the solar wind/IMF conditions. However, whether plasma sheet mesoscale disturbances are affected by the solar wind/IMF ULF fluctuations remains to be investigated.

In this study we investigate how plasma sheet electron temperature (T_e) observed by Time History of Events and Macroscale Interactions during Substorms (THEMIS) is affected by primary solar wind/IMF parameters, especially the effects from their ULF fluctuations, and by *AE* levels. We focus on T_e in the region around midnight geosynchronous distance ($6 \le r \le 8$ R_E and $21 \le \mathrm{MLT} \le 3$), where the earthward edge of the electron plasma sheet connects with the ring current. We investigate the correlations of observed T_e with the solar wind/IMF and *AE* parameters, as well as the correlations between these parameters. The THEMIS data selection and observational results are presented in section 2. To decouple the effects on T_e from different parameters, we conduct regression analysis by using a machine learning technique, support vector regression machine (SVRM). The principles of SVRM and how we construct SVRM models are described in section 3 and Appendix A. The SVRM model predictions of the effects from each solar wind parameter on T_e are presented in section 4 and possible physical explanations are discussed in section 5.

2. Observations

2.1. Data

In this study, we use 8 years (January 2008 to December 2015) of plasma and magnetic field measured by the three THEMIS probes (TH-A, TH-D, and TH-E). THEMIS provides particles and plasma moments measured by the electrostatic analyzer (ESA, 0.006–20 keV/q for ions and 0.007–26 keV for electrons) [*McFadden et al.*, 2008] and the solid state telescope (SST, 35 keV–6 MeV for ions and 30 keV–6 MeV for electrons). Full plasma distributions with time resolution of either ~1 or 6 min are used. Contamination has been removed by using the THEMIS data processing procedures. Pitch angle-averaged electron temperatures are obtained from the combined ESA and SST data. The spin-fit magnetic field vectors are measured by the FluxGate Magnetometer (FGM) instrument [*Auster et al.*, 2008].

We use 1 min solar wind parameters, including IMF B_y and IMF B_z in GSM coordinates, and the solar wind speed (V_{sw}) and solar wind density (N_{sw}). These solar wind data are measured by ACE, Wind, or Geotail. The arrival time of the IMF at the subsolar bow shock at ($x = 17$, $y = 0$, $z = 0$ R_E) is determined by calculating the minimum variance direction using the minimum variance analysis technique [*Weimer et al.*, 2003; *Weimer*, 2004]. To indicate the strength of ULF fluctuations of these parameters, we computed their ULF power parameters, including $B_{y,\,\mathrm{pow}}$ for IMF B_y, $B_{z,\,\mathrm{pow}}$ for IMF B_z, V_{pow} for V_{sw}, and N_{pow} for N_{sw}. We use the same methodology of computing ULF powers as that used in *Kim et al.* [2009] in the following steps: (1) fill the missing data points with a cubic spline interpolation; (2) subtract the background value by using a smoothing function with a window of 61 points (1 h); (3) use a fast Fourier transform (FFT) algorithm with a 128 point (2 h) moving window (2 min time shift between successive windows) to produce the power spectral density in the ULF Pc5 frequency range; here a Hanning window is used to minimize edge effects that result in spectral leakage in the FFT spectrum; and (4) compute the ULF power index, which in this study is defined as the sum of the power spectral densities over the frequencies 0.5–8.3 mHz. The units are nT^2 for $B_{y,\,\mathrm{pow}}$ and $B_{z,\,\mathrm{pow}}$, $\mathrm{km}^2/\mathrm{s}^2$ for V_{pow}, and cm^{-6} for N_{pow}. Figures 1a and 1b show two examples of IMF B_z with small and large $B_{z,\,\mathrm{pow}}$, respectively. Figures 1c and 1d show two examples of slow solar wind with small and large V_{pow}, respectively.

As shown by the previous studies described in section 1, the time-lag effect of the solar wind drivers on the magnetosphere is important, and the state of the magnetosphere at $t = t_0$ depends more on the solar wind

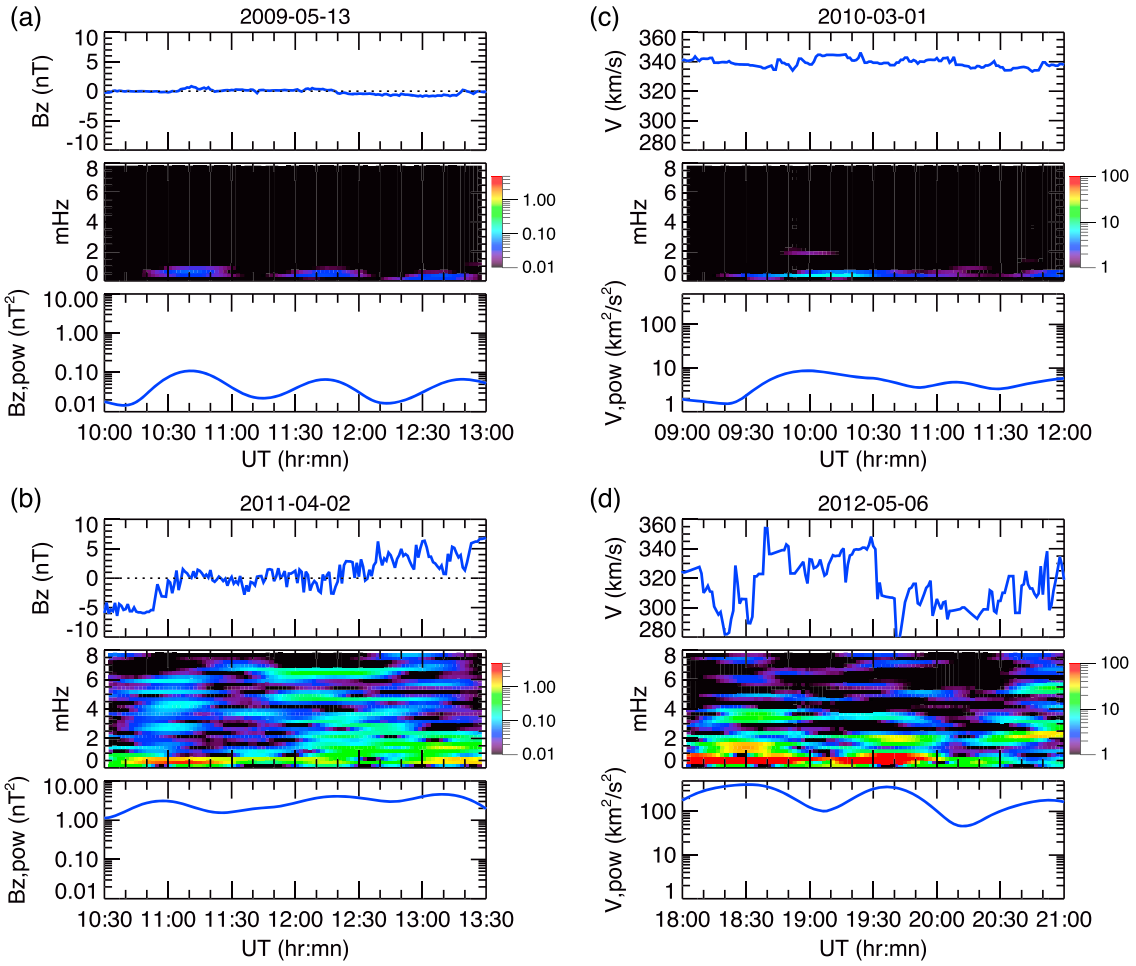

Figure 1. IMF B_z, its dynamic power spectrum, and ULF power for events of (a) low $B_{z,\mathrm{pow}}$ and (b) high $B_{z,\ \mathrm{pow}}$. The solar wind speed, its dynamic power spectrum, and ULF power for events of (c) low V_{pow} and (d) high V_{pow}.

history within an interval between $t = t_0$ and $t_0 - \Delta t$ than on the solar wind conditions at a single moment, $t = t_0 - \Delta t$. While there can be different ways to construct a parameter to represent the history of the solar wind conditions, in this study we computed time averages of the above solar wind parameters following the methodology of *Terasawa et al.* [1997]. For each data point at $t = t_0$ we compute time averages of each parameter over the time interval between $t = t_0$ and $t_0 - \Delta t$, where $\Delta t = 1$, 2, and 3 h. If the number of missing data points within Δt is more than 10% of the number of total points, then the average will not be used. Note that we have investigated the correlations of T_e with these solar wind parameters in their different hour averages from 1 to 8 h. We found that the correlation is the strongest with IMF B_z, and the IMF B_z correlation is the highest with its 3 h average. In addition, we found that it is sufficient to use averages up to 3 h to make SVRM prediction with satisfactory accuracy. Therefore, in this study we consider the solar wind history up to 3 h.

For this analysis, we selected THEMIS data points within $6 \le r \le 8\ R_E$, $21 \le \mathrm{MLT} \le 3$, and plasma beta > 0.01. The total number of available data points is 32,518. Each data point includes the information of T_e and its 30 corresponding parameters. These parameters are the three spatial parameters (r, MLT, and plasma beta), the eight solar wind parameters (IMF B_y, IMF B_z, V_{sw}, N_{sw}, $B_{y,\ \mathrm{pow}}$, $B_{z,\ \mathrm{pow}}$, V_{pow}, and N_{pow}) in their 1, 2, and 3 h averages, and the *AE* parameter in its 1, 2, and 3 h averages. Here plasma beta is used as a spatial parameter to indicate the z distance from the equatorial plane, since plasma beta at a fixed x-y location generally becomes lower with increasing z distances. Figure 2 shows the probability distributions of T_e, and the 1 h averages of the *AE* and solar wind parameters, as well as their quartile values. T_e in this region varies from ~0.1 to 10 keV with the median value of ~1.3 keV.

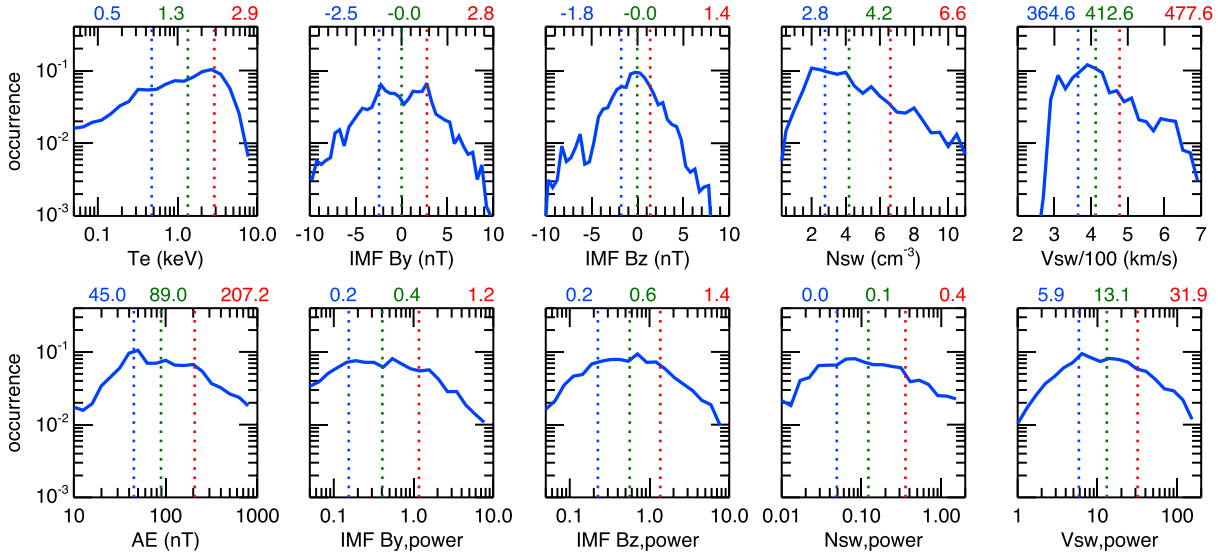

Figure 2. Probability distributions of plasma sheet electron temperature, solar wind/IMF parameters, and *AE*. The vertical dotted lines indicate the 25% (blue), 50% (green), and 75% (red) quartiles with their values indicated on the top of each plot.

2.2. T_e Variations With Locations, Solar Wind, Solar Wind ULF Powers, and *AE*

Figure 3a shows the T_e variations with the three spatial parameters and with the *AE* and solar wind parameters in their 3 h averages. Linear correlation coefficients between T_e and these parameters are indicated on the top of each plot. For the spatial dependences, there are no strong variations with r and MLT, but T_e decreases significantly with decreasing plasma beta. T_e increases with increasing *AE*, and their correlation coefficient is >0.5. Among the solar wind parameters, general increasing or decreasing trends are seen with IMF B_z, V_{sw}, N_{sw}, and V_{pow} and their correlation coefficients are >0.3. T_e, in general, becomes higher with increasing V_{sw} and V_{pow}, while it becomes lower with increasing IMF B_z and N_{sw}. The dependences on IMF B_z, N_{sw}, V_{sw}, and *AE* are consistent with the previous studies with ion temperatures [e.g., *Wang et al.*, 2007, 2009]. In comparison, variations with IMF B_y, $B_{y,\,pow}$, $B_{z,pow}$, and N_{pow} show no clear increasing or decreasing trends and their correlation coefficients are < ~0.2.

It is important to note that the T_e variation and the linear correlation with each of these parameters shown in Figure 3a are without restricting the values of the other parameters; therefore, the effects of the other parameters are not yet decoupled. Those parameters with relatively high correlations are very likely a dominant factor in affecting T_e. On the other hand, for a parameter with a very low correlation, the correlation may not necessarily indicate that it always has very weak effect on T_e. For example, the correlation coefficient between T_e and $B_{z,\,pow}$ is very low (0.14) under all conditions as shown in Figure 3a, but it can change significantly under different IMF B_z conditions. Figure 3b shows the correlations between T_e and $B_{z,\,pow}$ for three different IMF B_z ranges. It can be seen that the correlation is very low (0.11) during southward IMF, but it is high (0.6) during northward IMF. This shows that the effect of $B_{z,\,pow}$ on T_e is conditional depending on the conditions of other parameters. Therefore, to determine how the solar wind ULF fluctuations affect T_e, it is necessary to decouple their effects from other parameters.

2.3. Correlations Between Solar Wind and *AE* Parameters

Those solar wind parameters considered in this study are not entirely independent of each other. Figure 4 shows the correlations between different solar wind parameters in their 3 h averages, with their linear correlation coefficients indicated on the top of each plot. The 10%, 50%, and 90% percentiles are overplotted. IMF B_z is not correlated with either V_{sw} or N_{sw}, but conditions of high V_{sw} (> ~600 km/s) tends to occur at smaller IMF B_z magnitudes, while conditions of large N_{sw} (> ~10 cm^{-3}) tends to occur at larger IMF B_z magnitudes. These correlations are the same for IMF B_y (not shown). Also, IMF B_y is independent of IMF B_z (not shown).

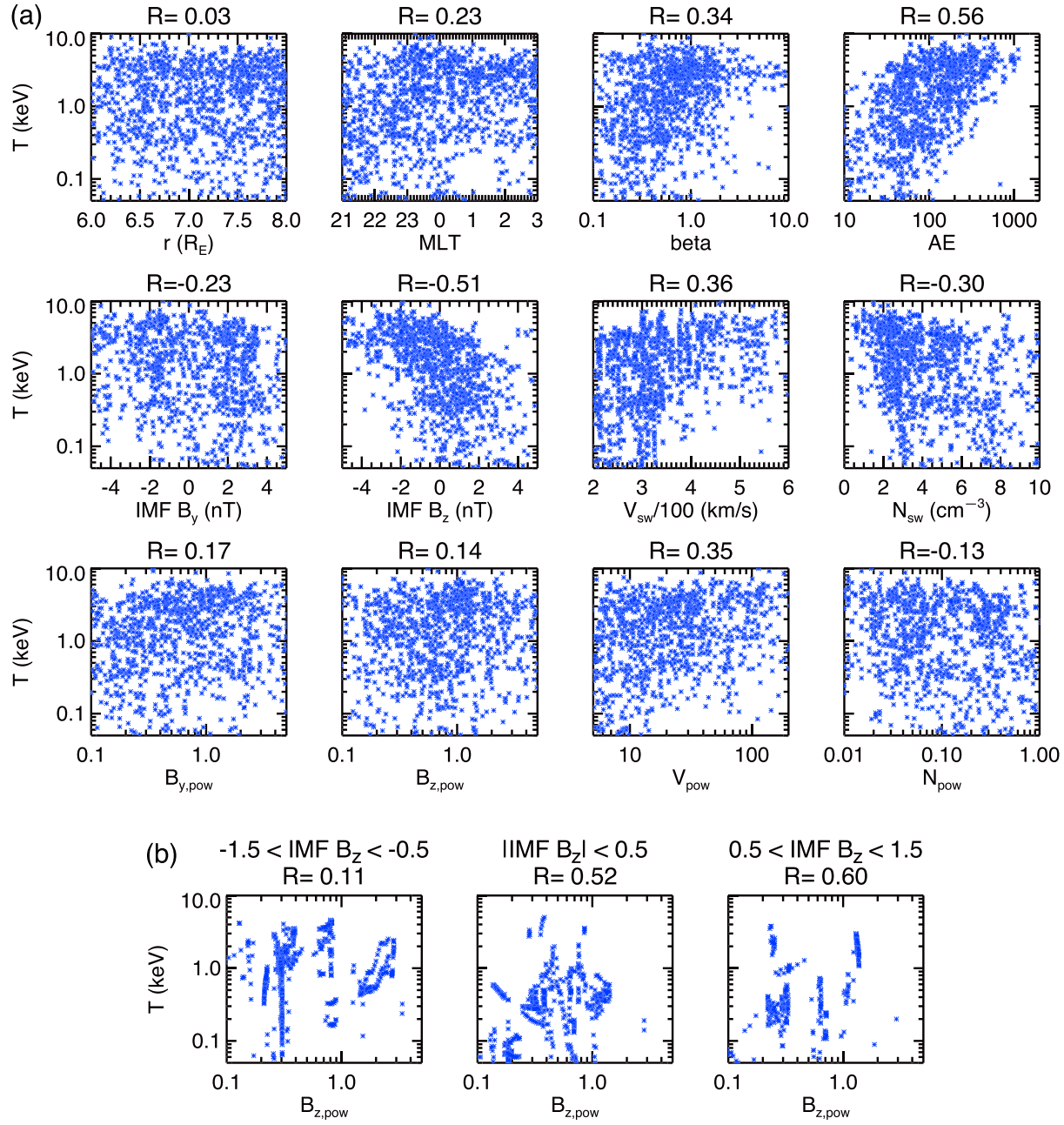

Figure 3. (a) T_e as a function of radial distances, MLTs, plasma beta, and *AE* and solar wind parameters (in their 3 h averages). The linear correlation coefficients are indicated on the top of each plot. (b) T_e as a function of $B_{z,\ pow}$ within different IMF B_z ranges.

V_{sw} and N_{sw} are anticorrelated, and variations of N_{sw} are much larger when V_{sw} is lower. For ULF powers, $B_{z,\ pow}$ is independent of IMF B_z, and it is stronger when V_{sw} becomes higher. V_{pow} is highly correlated with V_{sw}, while N_{pow} is highly correlated with N_{sw}. $B_{z,\ pow}$ and V_{pow} are clearly correlated, but V_{pow} and N_{pow} are independent of each other. The correlations of $B_{z,\ pow}$ with other solar wind parameters are consistent with the study of *Kessel* [2008].

Figure 4 also shows variations of *AE* with the solar wind parameters. *AE* is higher with decreasing IMF B_z and increasing V_{sw} but is independent of N_{sw}. Similarly, *AE* is higher with increasing $B_{z,\ pow}$ and V_{pow} but is almost independent of N_{pow}. This suggests that the solar wind ULF powers can affect plasma sheet mesoscale disturbances.

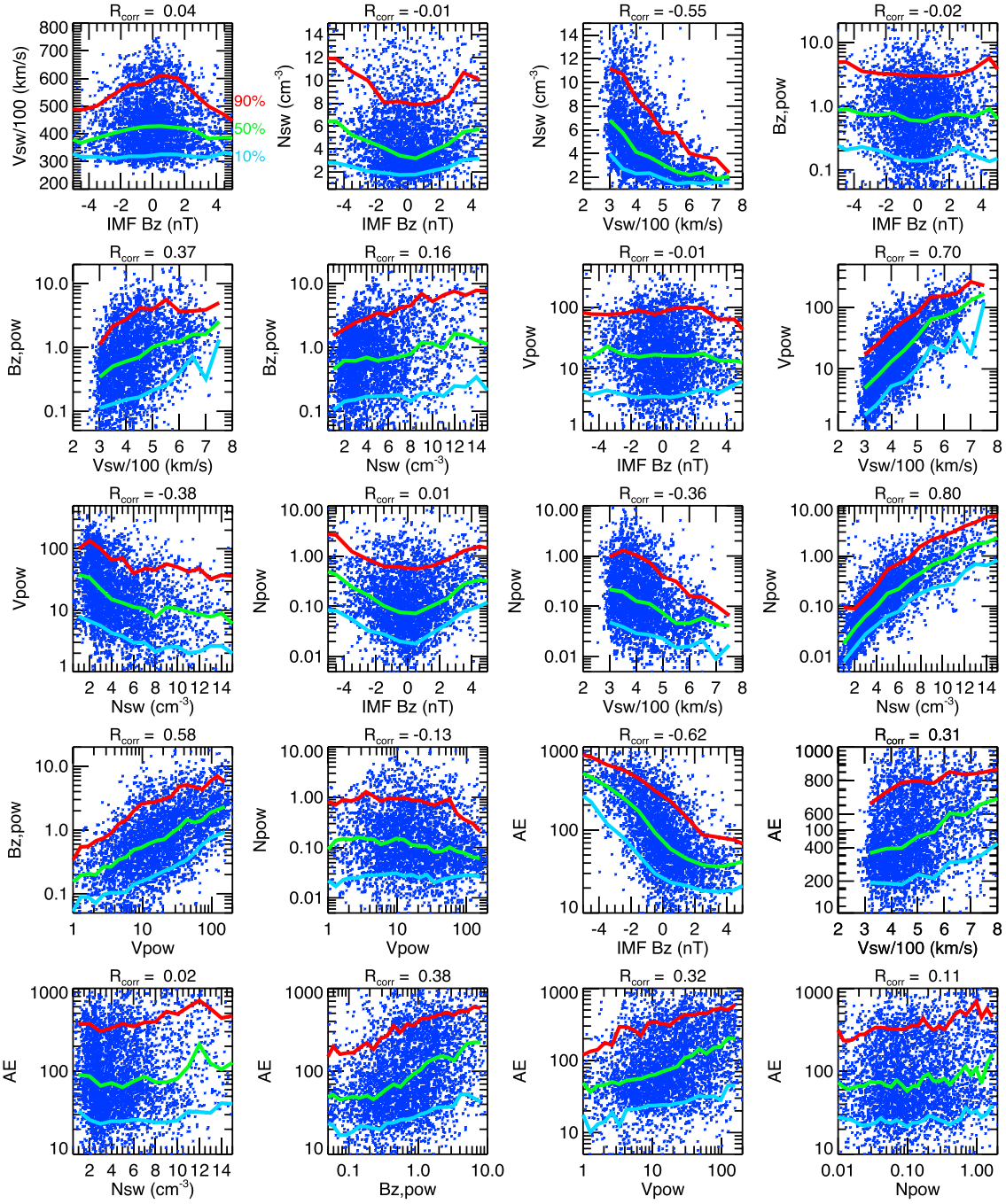

Figure 4. Correlations between different solar wind and *AE* parameters (in their 3 h averages). The blue, green, and red lines indicate the 10%, 50%, and 90% percentiles, respectively.

3. Regression Analysis

We conduct regression analysis to decouple the effects from different parameters in order to evaluate the contribution of solar wind ULF fluctuations on T_e. There are many different techniques available for such decoupling. In addition to traditional approach of empirical fitting observations with analytical formulas of the correlations, such as studies of *Tsyganenko and Mukai* [2003], *Nagata et al.* [2007], and *Dubyagin et al.* [2016], many new techniques based on machine learning have been developed, such as neural network and SVRM, and have been applied to understand the near-Earth magnetosphere [*Bortnik et al.*, 2016; *Yue et al.*, 2015a, 2015b]. Each of these techniques has its advantages and disadvantages. In the following we

describe the principles and advantages of SVRM and demonstrate the appropriateness of using SVRM for this study.

3.1. Support Vector Regression Machine

SVRM is regression analysis using the support vector machine (SVM) [*Vapnik and Lerner*, 1963]. SVM is a supervised nonlinear machine learning technique in the framework of statistical learning theory. The basic principle for a SVM model is to represent the training data set (training data are the data points used to train the model) as vectors in space mapped so that the data are divided to separate groups by a gap as wide as possible. New data from cross validation or a test data set are then mapped into the same space. Based on which side of the gap they fall, they are predicted to belong to one group or another. The SVRM maps multidimensional data (for example, in our case, locations and the solar wind parameters, these are called control parameters) into a high-dimensional feature space (in our case, electron temperature), via nonlinear mapping through a selected kernel function, and to perform a linear regression in this space. The key mathematical challenge of SVRM is to minimize the sum of the model errors for all the data points and a complexity term that enforces the flatness of the surface in the high-dimensional feature space. This leads to solving a quadratic programming problem, which is uniquely solvable [*Vapnik*, 1995]. The solution is represented by a subset of the training data as support vectors that are automatically selected by SVM/SVRM algorithm and can be used to uniquely determine the fitted surface. More comprehensive information of SVRM can be found in *Smola and Schölkopf* [2004].

We choose to use SVRM for this study because of its many advantages over other models: (1) SVRM is nonparametric regression so it requires no predetermined equation forms. (2) It is appropriate for control parameters that are likely to interact with each other (in our case, many of the solar wind parameters, as shown in Figure 4, correlate with each other). (3) Theoretically, it can bypass "the curse of dimensionality," which is the problem of available data becoming sparse when the dimensionality increases. (4) It avoids traps by local minima. In space physics, SVM has been used in forecasting, including solar $F_{10.7}$ [*Huang et al.*, 2009], solar wind velocity [*Liu et al.*, 2011], substorms [*Gavrishchaka and Ganguli*, 2001], *Kp* [*Ji et al.*, 2013], and storm time ionosphere [*Sun et al.*, 2011]. SVRM has been applied to model the magnetopause shape [*Wang et al.*, 2013]. Recently, *Yue et al.* [2015a, 2015b] have successfully applied SVRM to evaluate how the observed plasma sheet plasma pressure is controlled by solar wind energy loading and *AE*.

To demonstrate that SVRM is capable of decoupling the effects on a quantity from different control parameters similar to those of this study, we have tested SVRM analysis and prediction with a virtual data set of plasma sheet ion temperatures prescribed by an analytical coupling function of spatial and solar wind parameters given by the *Tsyganenko and Mukai*, 2003 (TM03) model [*Tsyganenko and Mukai*, 2003]. This test, as well as the procedures of constructing a SVRM model, is presented in Appendix A.

3.2. SVRM T_e Model

To construct a SVRM model to predict T_e following the procedures described in Appendix A, we use the radial basis function combined with ε-SVR ($\varepsilon = 0.1$, $C = 1$, and $\gamma = 1$) and a training data from 30% of our data set after experimenting with different functions, SVR types, and values of these free parameters. We have also experimented with different sets of control parameters chosen from the 30 parameters described in section 2.1, which include the three spatial parameters, the eight solar wind parameters in their three time averages, and the *AE* parameter in its three time averages. Figures 5a–5e show the testing results from the models using five different parameter sets: (1) using all of the 30 parameters; (2) excluding all IMF B_y, $B_{y,\mathrm{pow}}$, and *AE* parameters from the 30 parameters; (3) the same as (2) but use only the 1 h averages; (4) excluding all IMF B_y, *AE*, and all ULF power parameters from the 30 parameters; and (5) using *AE* parameters only. Comparing set 1 with set 2 indicates that it is not essential to use IMF B_y, $B_{y,\,\mathrm{pow}}$, and *AE* as control parameters if other solar wind parameters are included. This suggests that the effects of IMF B_y and $B_{y,\,\mathrm{pow}}$ on T_e are much weaker than other parameters. As shown in Figure 4, *AE* is closely correlated with the solar wind parameters. We have tested using SVRM to predict *AE* with these solar wind parameters (excluding IMF B_y and $B_{y,\mathrm{pow}}$) and found that *AE* can be well predicted, as indicated by the high correlation between observed *AE* and model *AE* shown in Figure 5f. This suggests that the *AE* parameter becomes redundant in the presence of the solar wind parameters and thus needs not to be included. Comparing set 2 and 3 indicates that it is important to consider a longer history of solar wind conditions. Note that we have tested including longer time

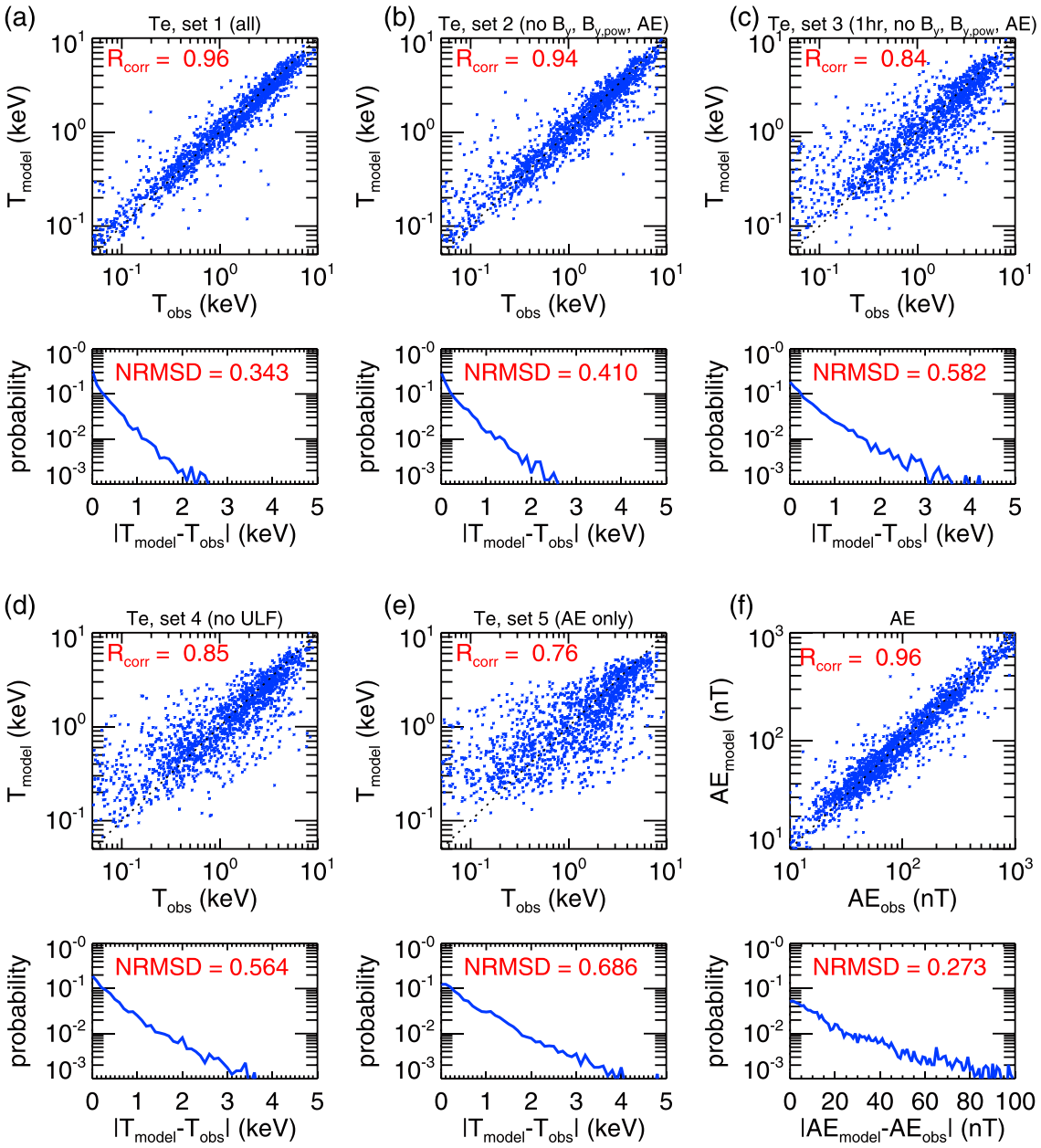

Figure 5. (a–e) Observed T_e versus model T_e (the linear correlation coefficient indicated in red) (top) and probability distributions of the differences between observed T_e and model T_e (NRMSD indicated in red) (bottom) for different control parameter sets. (f) Comparisons between observed *AE* and model *AE*.

averages from 4 to 8 h, but the prediction results are only slightly better. The differences between set 3 and 4 indicate that ULF powers are important to T_e. Despite that Figure 3a shows that *AE*, among all the parameters, has the highest linear correlation with T_e, Figure 5e shows that it is far from sufficient to predict T_e using only the *AE* parameters. Based on these comparisons, in section 4 we present the prediction results from the SVRM T_e model constructed from the control parameter set 2.

4. Model Results

In this section we present the SVRM predictions of the T_e variations with IMF B_z, V_{sw}, and N_{sw} when their ULF powers are weak, as well as the T_e variations with $B_{z,\,pow}$, V_{pow}, and N_{pow}.

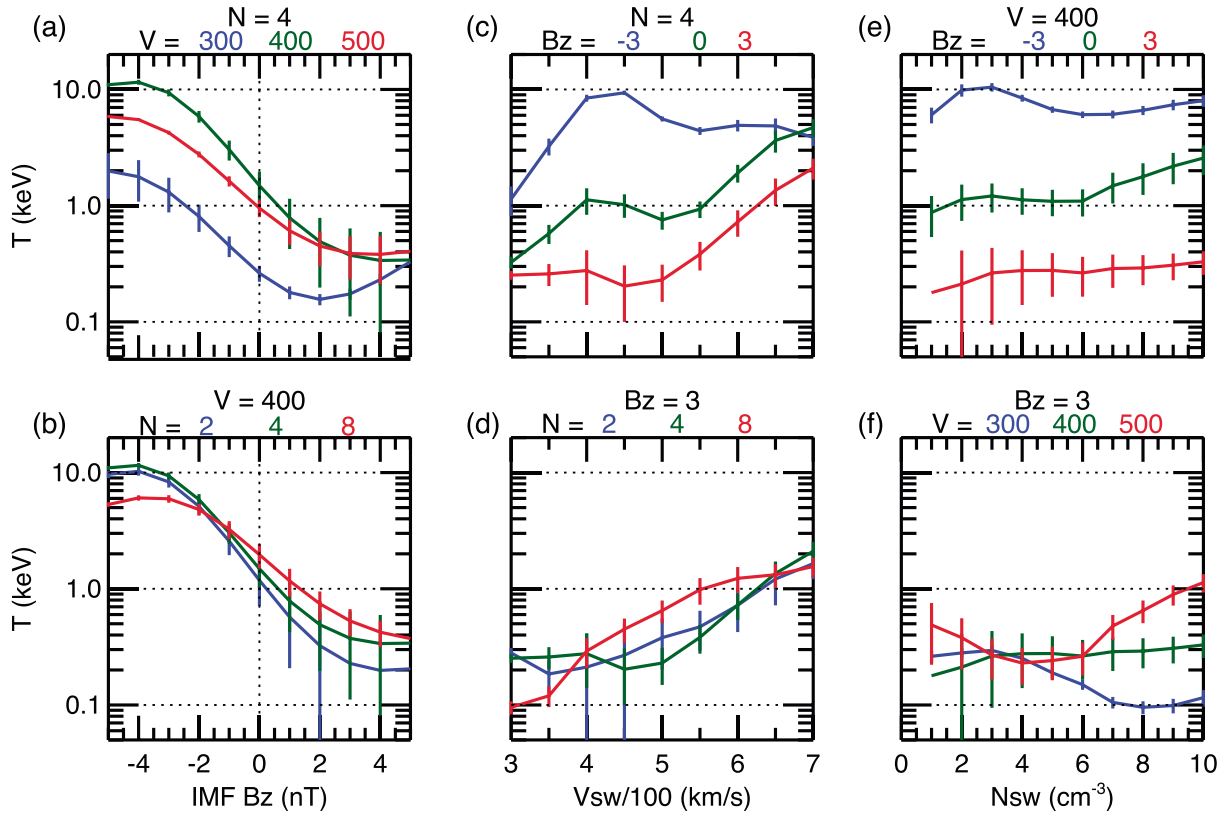

Figure 6. SVRM T_e as a function of IMF B_z for (a) $N_{sw} = 4\ cm^{-3}$ and (b) $V_{sw} = 400$ km/s, as a function of of V_{sw} for (c) $N_{sw} = 4\ cm^{-3}$ and (d) IMF $B_z = 3$ nT, and as a function of N_{sw} for (e) $V_{sw} = 400$ km/s and (f) IMF $B_z = 3$ nT. The different line colors indicate values of the parameter in comparisons. For all plots, $B_{z,\ pow} = 0.1$ is specified, and the values of the 10% percentiles listed in Table 1 are specified for V_{pow} and N_{pow}.

4.1. Dependences on IMF B_z, V_{sw}, and N_{sw}

We investigate in Figure 6 the variations of T_e with IMF B_z, V_{sw}, and N_{sw} when the solar wind ULF powers are low. To specify this low ULF power condition, we use the 10% percentile values for $B_{z,\ pow}$, V_{pow}, and N_{pow} as shown in Figure 4. The 10% $B_{z,\ pow}$ is independent of IMF B_z, but it increases from ~0.1 to 0.3 as V_{sw} increases from 300 to 700 km/s. We thus specify a constant $B_{z,\ pow} = 0.2$. For the 10% percentile V_{pow} (N_{pow}), their values as a function of V_{sw} (N_{sw}) are listed in Table 1. We assume a steady condition so that the same values are specified for the1 h, 2 h, and 3 h averages. The model T_e is predicted at the location of $r = 7\ R_E$, MLT = 0, and plasma beta = 0.5. Note that the errors shown in Figure 6 are predicted by the T_e error model (see Appendix A). Figure 6a shows the T_e variations with IMF B_z with constant $N_{sw} = 4\ cm^{-3}$ (the median N_{sw} as shown in Figure 2) under three solar wind speeds as indicated by lines of different colors. Similarly, Figure 6b shows T_e variations with IMF B_z with constant $V_{sw} = 400$ km/s (the median V_{sw} as shown in Figure 2) under three solar wind densities. These variations show that T_e is lower during northward IMF than southward IMF. During southward IMF, T_e increases monotonically with increasing |IMF B_z| with the increase rates depending on the V_{sw} and N_{sw} conditions. For example, Figure 6a shows that T_e can increase by a factor of ~10 as IMF B_z changes from 0 to −5 nT when $V_{sw} = 300$ km/s (blue line), but Figure 6b shows that T_e

Table 1. 10%, 50%, and 90% Percentiles of V_{pow} (N_{pow}) as a Function of V_{SW} (N_{SW})

V_{SW}[a] (N_{SW})	300 (1)	350 (2)	400 (3)	450 (4)	500 (5)	550 (6)	600 (7)	650 (8)	700 (9)	750 (10)
10% V_{pow} (N_{pow})	1.7 (0.008)	2.7 (0.015)	5.2 (0.028)	6.2 (0.042)	10 (0.064)	23 (0.079)	20 (0.13)	40 (0.18)	17 (0.23)	121 (0.26)
50% V_{pow} (N_{pow})	4.8 (0.018)	7.8 (0.036)	13 (0.064)	20 (0.1)	34 (0.18)	64 (0.22)	72 (0.34)	88 (0.45)	130 (0.65)	166(0.81)
90% V_{pow} (N_{pow})	17.2 (0.09)	25 (0.09)	42 (0.148)	61 (0.26)	83 (0.54)	149 (0.7)	153 (0.94)	174 (1.56)	260 (2.1)	227 (2.4)

[a] V_{SW} in km/s, N_{SW} in cm^{-3}, V_{pow} in km^2/s^2, and N_{pow} in cm^{-6}.

increases only by a factor of 2 when $N_{sw} = 8\,cm^{-3}$ (red line). In comparison with southward IMF, T_e variations with |IMF B_z| during northward IMF are smaller. These predicted T_e variations with IMF B_z are consistent with the conclusions from the previous observational studies described in section 1.

Figure 6c shows that T_e variations with V_{sw} are very different under different IMF B_z conditions. During northward IMF, T_e increases significantly with increasing V_{sw}. For example, for IMF $B_z = 3$ nT (red line), T_e increases by a factor of ~10 as V_{sw} increases from 300 to 700 km/s. In contrast, for southward IMF (blue line), T_e increases significantly from $V_{sw} = 300$ to 400 km/s, but it becomes almost independent of V_{sw} for $V_{sw} \geq$ ~500 km/s. Figure 6d shows that the T_e variations with V_{sw} during northward IMF are similar under different N_{sw} conditions.

In comparison with the T_e variations with IMF B_z and V_{sw}, Figures 6e and 6f show that T_e variations with N_{sw} are much weaker. As shown in Figure 6e, when $V_{sw} = 400$ km/s, T_e is almost independent of N_{sw} under all IMF B_z conditions. When N_{sw} is high, Figure 6f shows that T_e increases with increasing N_{sw} under higher V_{sw} (red line) but decreases under lower V_{sw} (blue line).

4.2. Dependences on Solar Wind ULF Powers

Figure 7 shows the T_e variations with $B_{z,\,pow}$ under different IMF B_z (different line colors), V_{sw} (different columns), and N_{sw} (different rows) conditions. Similar to Figure 6, the 10% percentile V_{pow} (N_{pow}) values as a function of V_{sw} (N_{sw}) are specified. The variation trends are very different depending on IMF B_z; for northward IMF (red lines), T_e increases with increasing $B_{z,\,pow}$ under most of the conditions, except for $V_{sw} = 400$ km/s and $N_{sw} = 6\,cm^{-3}$. The increase tends to be more significant under higher V_{sw}. For example, for $N_{sw} = 2\,cm^{-3}$, T_e increases by a factor of ~10 as $B_{z,\,pow}$ increases from 0.3 to 5 when V_{sw} is 500 km/s, but increases by a smaller factor of ~3 when V_{sw} is 300 km/s. For southward IMF (blue lines), T_e is almost independent of $B_{z,\,pow}$ when V_{sw} is low (for example, 300 km/s). But under higher V_{sw} (for example, $>$ 400 km/s), as $B_{z,\,pow}$ increases T_e first remains similar or increases slightly then decreases. For IMF $B_z = 0$ (green lines), T_e, in general, increases with increasing $B_{z,\,pow}$ and the increasing rate becomes more significant when N_{sw} is higher. The fact that the predicted T_e variations are large with $B_{z,\,pow}$ during weak IMF B_z and northward IMF but small during southward IMF is in agreement with the observed variations shown in Figure 3b.

We investigate in Figure 8a the effect of V_{pow} on T_e for southward IMF (top row), IMF $B_z = 0$ (middle row), and northward IMF (bottom row), and for $N_{sw} = 2$ (left column) and $4\,cm^{-3}$ (right column). Note that the T_e scales for different IMF B_z conditions are different in order to show the trends more clearly. Constant $B_{z,\,pow} = 0.2$ is specified. Since the ranges of V_{pow} change with V_{sw}, we plot T_e as a function of V_{sw} and compare temperatures corresponding to low V_{pow} (10% percentile, blue lines), median V_{pow} (50% percentile, green lines), and high V_{pow} (90% percentile, red lines). Like the T_e dependence on V_{sw} as shown in Figure 6c, Figure 8a shows that T_e dependence on V_{pow} is also quite different under different IMF B_z conditions. For the same V_{sw}, in general, temperature is higher when V_{pow} is higher under northward IMF and IMF $B_z = 0$, but this correlation is opposite under southward IMF. The V_{pow} is more effective for $V_{sw} \leq$ ~500 km/s and during northward IMF. For example, for IMF $B_z = 3$ nT and $N_{sw} = 4\,cm^{-3}$, T_e at $V_{sw} = 450$ km/s is a factor of ~6 higher for the high V_{pow} than for the low V_{pow}. Figure 8b shows the T_e variations with N_{sw} for low, median, and high N_{pow} (different color) for southward (top row), IMF $B_z = 0$ (middle row), and northward IMF (bottom row), and for $V_{sw} = 300$ (left column) and 400 km/s (right column). The effect of N_{pow} is mainly seen during IMF $B_z = 0$ and northward IMF for $N_{sw} \geq 5\,cm^{-3}$. And for these favorable IMF and N_{sw} conditions, T_e for the same N_{sw} is higher when N_{pow} is higher. For example, for IMF $B_z = 3$ nT and $V_{SW} = 300$ km/s, T_e at $N_{sw} = 8\,cm^{-3}$ is a factor of ~5 higher for high N_{pow} than for low N_{pow}.

The above results shown in Figures 6–8 indicate that the solar wind ULF fluctuations have a substantial effect on T_e during weak IMF B_z or northward IMF. Considering the T_e occurrence distributions shown in Figure 2, when $B_{z,\,pow}$ is low ($\lesssim$0.2), the high T_e ($>$3 keV, the 75% quartile) can only occur under the condition of strong southward IMF ($\lesssim -3$ nT) and high V_{sw} ($\gtrsim$400 km/s). However, for the same high V_{sw} ($\gtrsim$400 km/s), even when IMF B_z is very weak (~0), the high T_e can still occur if $B_{z,\,pow}$ is high ($\gtrsim$2). On the other hand, while T_e becomes lower when northward IMF is stronger and V_{sw} is lower, the low T_e ($<$0.5 keV, the 25% quartile) only occurs when none of $B_{z,\,pow}$, V_{pow}, and N_{pow} are high.

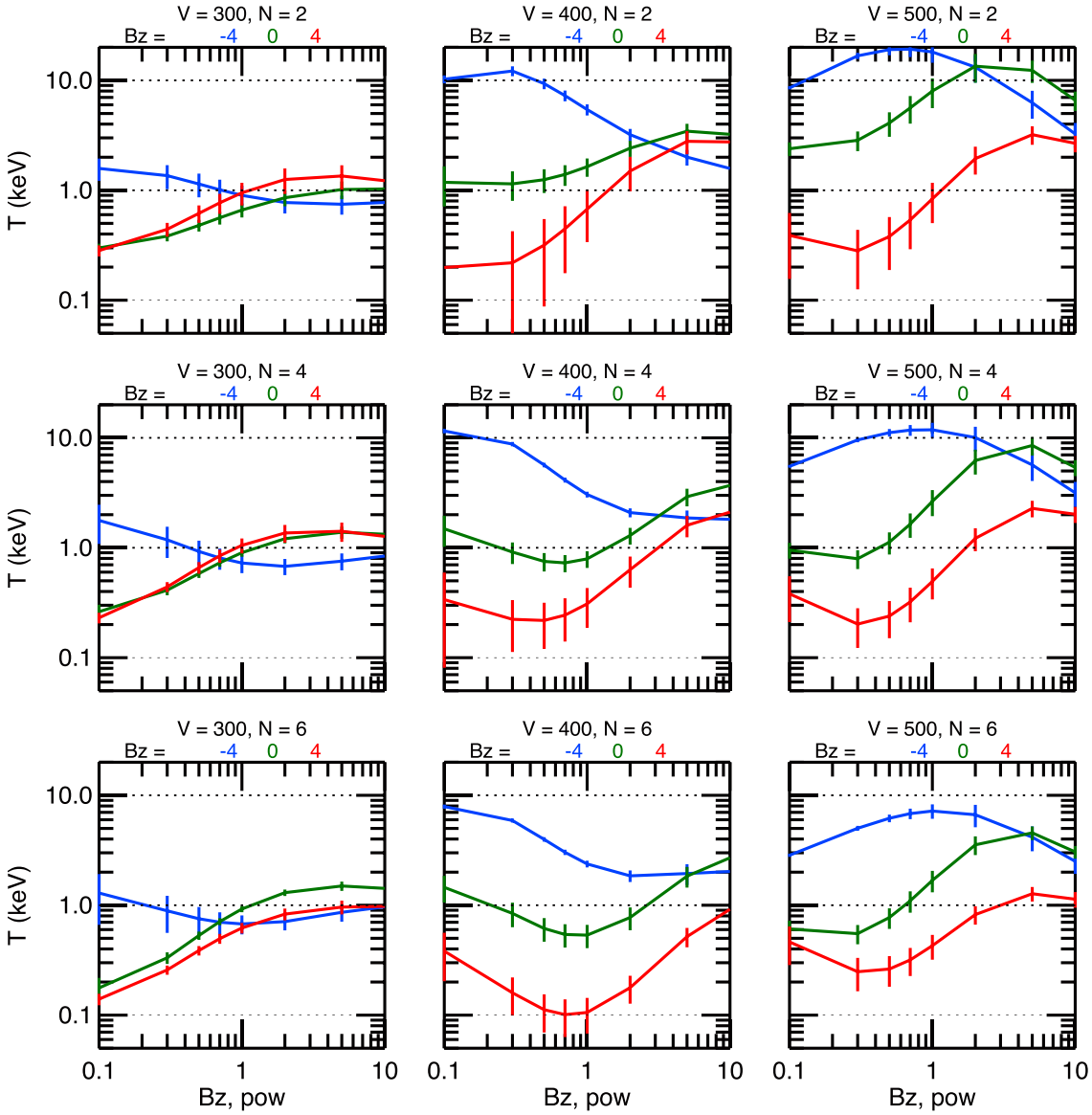

Figure 7. SVRM T_e as a function of $B_{z,\ pow}$ for IMF $B_z = -4$ (blue lines), 0 (green lines), and 4 nT (red lines) under different V_{sw} (different columns) and N_{sw} (different rows) conditions. The values of the 10% percentiles listed in Table 1 are specified for V_{pow} and N_{pow}.

5. Discussion

As shown in Figures 6 and 7, high T_e near midnight geosynchronous distance can occur under either of the two IMF conditions: (1) strong southward IMF B_z with weak $B_{z,\ pow}$ or (2) weak IMF B_z or northward IMF with strong $B_{z,\ pow}$. Here we discuss possible large-scale and mesoscale processes that could enhance T_e through inward transport and energization of high-energy electrons (above the plasma sheet thermal energy) under these two conditions. Note that these processes may also lead to generation of many different small-scale waves; however, complicated electron energization by these waves remains to be better understood and thus is not discussed here. For high-energy plasma sheet electrons, their drift paths and Alfvén layers are determined by the competition between sunward electric drift and azimuthal magnetic drift. Their Alfvén layers around midnight during low *AE* (<50 nT) are located at $r \sim 7.5$–$8\ R_E$ [*Jiang et al.*, 2011]. Under the first IMF condition, enhanced electric drift under stronger southward IMF B_z drives Alfvén layers further inward [*Korth et al.*, 1999; *Wang et al.*, 2008, 2011], which can account for the T_e increase with increasing southward IMF B_z shown in Figures 6a and 6b.

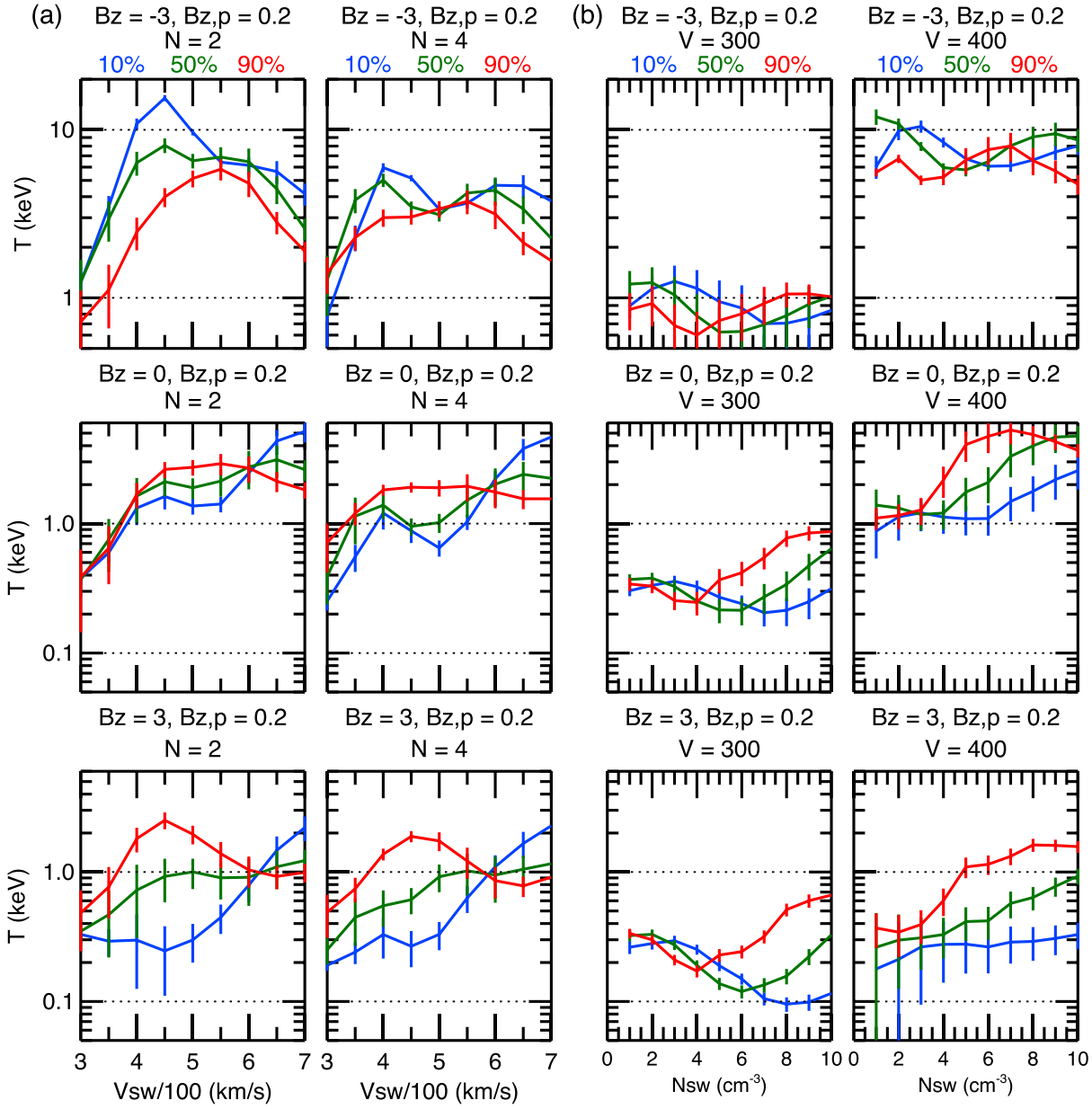

Figure 8. (a) SVRM T_e as a function of V_{sw} with V_{pow} in its 10% (blue), 50% (green), and 90% (red) percentiles listed in Table 1. (b) SVRM T_e as a function of N_{sw} with N_{pow} in its 10% (blue), 50% (green), and 90% (red) percentiles listed in Table 1. The values specified for IMF B_z, $B_{z,\,pow}$, and N_{sw} are indicated on the top of each plot.

On the other hand, under the second IMF condition of weak IMF B_z or northward IMF, large-scale convection is not sufficiently strong to bring high-energy electrons to geosynchronous distance. But under the condition of higher $B_{z,\,pow}$ when plasma sheet mesoscale disturbances are expected to be stronger, as suggested by the correlation between *AE* and $B_{z,\,pow}$ shown in Figure 4, high-energy electrons could get access to smaller radial distances through other processes. Particle simulations have shown that strong transient electric field associated with an earthward-moving dipolarization front (a plasma bubble) could locally inject energetic particles to the inner magnetosphere [e.g., *Gabrielse et al.*, 2016; *Yang et al.*, 2016]. In addition, the magnetosphere-ionosphere coupling through field-aligned currents associated with these mesoscale perturbations results in perturbations in the convection electric field [e.g., *Gkioulidou et al.*, 2009]. These mesoscale electric field perturbations could violate the third adiabatic invariant and result in radial diffusion. *Liu et al.* [2003] considered continuous mesoscale electric field perturbations in

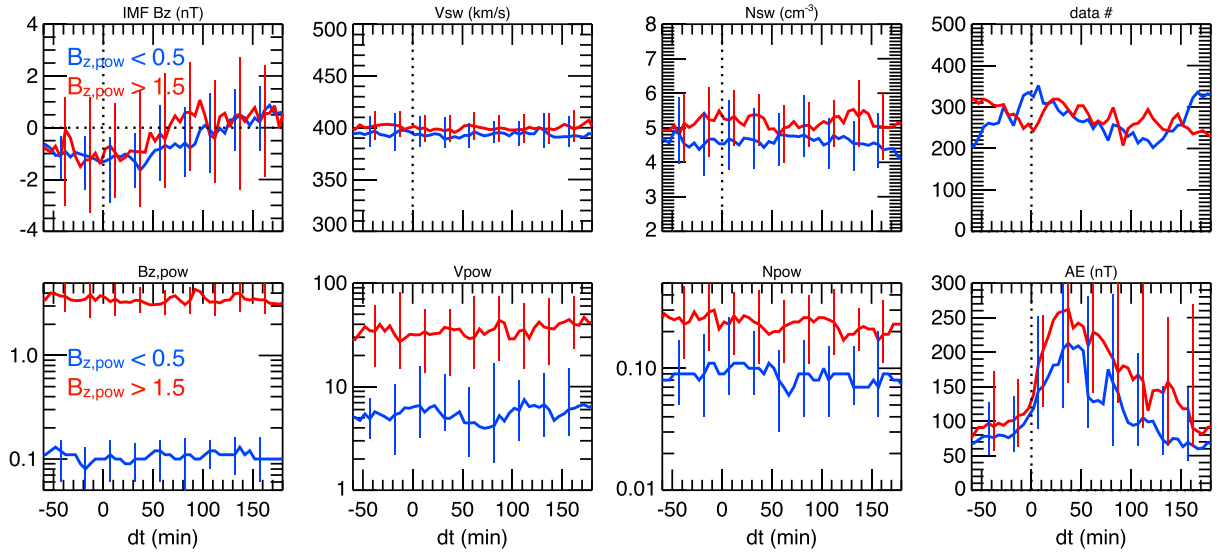

Figure 9. Medians of IMF B_z, the solar wind speed, the solar wind density, and their ULF powers, number of data points, and *AE* as a function of time relative to substorm onset at $t = 0$ for $B_{z,\ pow} < 0.5$ (blue) and $B_{z,pow} > 1.5$ (red) (the vertical lines indicate the 25% and 75% quartiles).

their kinetic simulations and showed that the resulting radial diffusion could bring high-energy (~150 keV) plasma sheet electrons ~2 R_E further inward in a time scale of a few hours in comparison with the results without such mesoscale perturbations. Therefore, considering continuous occurrence of mesoscale disturbances over several hours, such as during the substorm expansion and recovery phases and stronger disturbances under higher $B_{z,\ pow}$, it is plausible that inward radial diffusion can contribute to substantial T_e increase at geosynchronous orbits if $B_{z,\ pow}$ remains strong for a few hours.

To further investigate whether mesoscale disturbances during substorms are affected by $B_{z,\ pow}$, we analyzed statistically the variations of *AE* from the growth phase to recovery phases of isolated substorms. We selected 11 years of substorms from 1995 to 2005 from a substorm list composed by *Hsu and McPherron* [2012]. The onset of each substorm was identified by a sharp drop in *AL* of more than 100 nT in 20 min. We define an isolated substorm as the separation between its onset and the onsets of the two neighboring substorms being larger than 5 h. *AE* and solar wind/IMF data in 5 min resolutions from 1 h before to 3 h after the substorm onset at $t = 0$ were used. We restricted the solar wind conditions to $400 < V_{sw} < 450$ km/s and $3 < N_{sw} < 7\ \mathrm{cm}^{-3}$ (around their median values as shown in Figure 2) and separated the data points to correspond to low and high ULF power with $B_{z,\ pow} < 0.5$ and $B_{z,\ pow} > 1.5$, respectively. Temporal variations of the quartiles of different solar wind parameters and *AE* are shown in Figure 9. While the variations of IMF B_z, V_{sw}, and N_{sw} are similar for the low and high $B_{z,\ pow}$, *AE* is higher for higher $B_{z,\ pow}$ during the substorm expansion and recovery phases. Note also that IMF B_z during these two phases is weak or positive. These results thus suggest that stronger inward radial diffusion of high-energy electrons could possibly contribute to the high T_e under the second IMF condition. Since diffusion is a slower process than the drifts, this may also explain why the SVRM prediction is more accurate, as shown in Figure 5, when considering a longer history of the solar wind and IMF conditions.

The inward radial diffusion might also explain the V_{sw} dependences shown in Figure 6c that during northward IMF T_e becomes higher with increasing V_{sw} even when the solar wind ULF fluctuations are weak. Numerous simulations [e.g., *Nykyri*, 2013] and observations [e.g., *Taylor et al.*, 2012; *Lin et al.*, 2014] have shown that the Kelvin-Helmholtz (K-H) surface waves at the flank magnetopause occur more frequently under higher V_{sw} during northward IMF and that they are capable of driving Pc5 ULF magnetic field fluctuations at geosynchronous distance [*Rae et al.*, 2005; *Agapitov et al.*, 2009]. Therefore, the K-H surface waves could possibly result in Alfvén layer fluctuations and thus radial diffusion of high-energy plasma sheet electrons to geosynchronous distance even when there are no ULF fluctuations in the solar wind.

6. Summary

Using a machine learning technique, SVRM, we conduct regression analysis of the dependences of T_e in the region of $r = 6–8\ R_E$ and MLT = 21–3 on IMF B_z, V_{sw}, and N_{sw}, and their ULF fluctuation powers, $B_{z,\ pow}$, V_{pow}, and N_{pow}. Our main findings are the following:

1. T_e in this region can vary by 2 orders of magnitude from 0.1 to 10 keV. The median T_e is 1.3 keV, and the 25% and 75% quartiles are 0.2 and 3 keV, respectively.
2. When the solar wind ULF fluctuations are weak, among the three solar wind parameters IMF B_z has the strongest effect while N_{sw} has the least effect. T_e increases monotonically with increasing southward IMF. As V_{sw} increases, T_e, in general, increases but the rate of increase depends on the IMF B_z conditions. T_e can increase to >3 keV (above the 75% quartile) during strong southward IMF when V_{sw} is high. This T_e increase is likely due to inward transport of plasma sheet electrons to geosynchronous distance by stronger convection electric drift under this solar wind/IMF condition.
3. T_e can also increase to >3 keV even during weak or positive IMF B_z if $B_{z,\ pow}$ remains high for a few hours. Large-scale earthward electric drift under this IMF B_z condition is likely too weak to cause such increase. This T_e increase may be attributed by radial diffusion due to mesoscale electric drift perturbations, which can persist for a few hours during substorm recovery phase and are more enhanced under higher $B_{z,\ pow}$.

Appendix A

To demonstrate that SVRM has the capability of decoupling the effects on a plasma sheet property from multiple parameters, including spatial parameters and solar wind parameters, we evaluate the SVRM analysis and predictions with a virtual data set of ion temperature established from a known coupling function given by the TM03 model. The TM03 analytical coupling function, which is based on Geotail observations, specifies plasma sheet ion temperatures in the region of $r = 10–50\ R_E$ from four parameters: r, MLT, IMF B_z, and V_{sw} (see equation (4) of *Tsyganenko and Mukai* [2003]). We compose a virtual data set that consists of 30,000 data points (similar to the number of the THEMIS data set described in section 2.1) of the TM03 ion temperatures corresponding to spatial parameters randomly selected from $r = 10–20\ R_E$ and MLT = 21–0 and the IMF B_z and V_{sw} parameters randomly selected from the probability distributions shown in Figure 2. Temperature noise within ±10% was then randomly added to these TM03 temperatures. Variations of the virtual temperature with each of the four parameters are shown in Figure A1a without restricting the parameter ranges. These variations show that the virtual temperature increases substantially with decreasing r and IMF B_z and with increasing V_{sw}. In comparison, it appears almost independent of MLT.

A SVRM model is constructed with the following procedures:

1. Specifying kernel functions and SVM types: The kernel function is used to map the training data (see procedure (3) below) to high-dimensional feature space. Then in this feature space, SVM is used to obtain a smooth model surface. There are several functions available, such as the radial basis function (RBF), a linear function, a polynomial function, and a sigmoid function. There are several SVM types available for regression, such as ε-support vector regression (SVR) [*Vapnik*, 1995], ν-SVR [*Scholkopf et al.*, 2000], and least squares SVR. After experimenting with different kernel functions and SVM types, for this study we chose to use a combination of RBF and ε-SVR. This combination is commonly used and is very flexible and robust in handling many problems and is free from many difficulties of some other combinations. RBF is $\exp(-\gamma|u-v|^2)$, where u is data, v is prediction, and γ controls how far a discrete data point is effective in constructing a continuous fitted parameter function for fitting in the high-dimensional feature space. The ε in ε-SVR is the permissible error for a linear regression function.
2. Specifying fitting parameters: To use RBF and ε-SVR, there are three free parameters (ε, C, and γ) to be specified in the model fitting process. The ε, introduced through ε-SVR, is the distance from the fitted surface within which no fitting error is counted for a data point. The error factor, C, introduced through SVR, is defined for data points falling outside of distance ε. The γ is the parameter in RBF. Using different C and γ, SVRM gives different model distributions. A larger γ leads to a smaller data effective range, and model reflects more local observations. And a larger C leads to tighter fitting of the data points, which can adversely affect model smoothness (see examples shown in Figure 3 of *Wang et al.* [2013] and

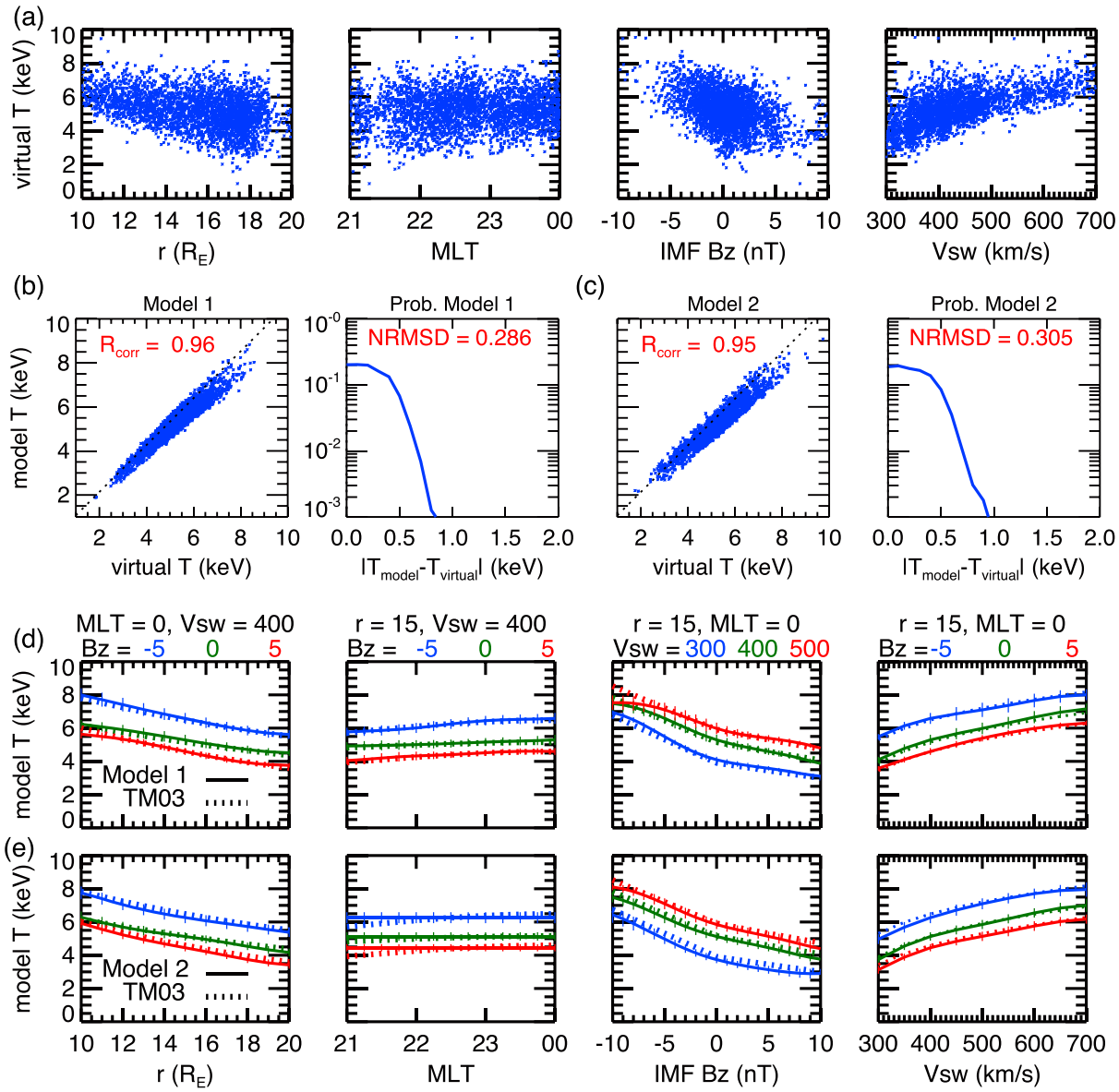

Figure A1. (a) Variations of the virtual T with r, MLT, IMF B_z, and V_{sw}. Virtual T versus model T (the liner correlation coefficients indicated in red) and probability distributions of the differences between virtual T and model T (NRMSD indicated in red) from (c) model 1 and (d) model 2. SVRM predictions (solid lines) from (d) model 1 and (e) model 2 in comparison with TM03 model (dashed lines). The error bars are estimated by using the error models.

Figure 6 of *Yue et al.* [2015a]). A grid search of these parameters with cross validation is used for determining their optimized values. For this study, we set $\varepsilon = 0.1$ and chose $\gamma = 1$ and $C = 1$ for smoothness.

3. Model training: To construct a model, the input data for training SVRM (training data set) include the quantity to be predicted and its control parameters. For example, in this virtual data set, ion temperature is the quantity and the control parameters can be any combination selected from the four parameters: r, MLT, IMF B_z, and V_{sw}. SVRM only needs a portion of a data set as training data. A smaller data percentage may not be sufficient for a model with satisfactory prediction, while larger percentage may take a longer computer processing time without really improving the prediction.
4. Model testing: To evaluate the model prediction, we use a testing data set which we randomly selected a portion (in this example, 30%) of the virtual data points (excluding the training data). We use the control parameters of these testing data points to predict temperature, T_{model}. We computed two prediction quantities: the linear correlation coefficient (R_{corr}) between T_{model} and the virtual temperatures, $T_{virtual}$,

and normalized root-mean-square deviation (NRMSD). NRMSD is defined as $\sqrt{\left(\sum (T_{\text{model}} - T_{\text{virtual}})^2\right)/n}/\sigma$, where n is the number of points and σ is the standard deviation of T_{virtual}.

We constructed several models by using different training data sets, which we used r, MLT, IMF B_z, and V_{sw} as the control parameters and different percentages (5% to 70%) of the virtual data set. Comparing the R_{corr} and NRMSD from different models show that using 30% of the data set as the training data is sufficient. Figure A1 b shows the correlations between T_{virtual} and T_{model} and the probability distribution of their differences for the model using 30% of the data set (model 1). To evaluate whether the SVRM model can recover the true dependences prescribed by the TM03 coupling function, we plot in Figure A1d variations of T_{model} with one of the four parameters while keeping the other parameters constant (solid lines). These model variations are in a very good agreement with the true variations from the TM03 (dashed lines), indicating that SVRM can reliably decouple the effects from different parameters. Note that the error bars shown in Figure A1d are predicted by a separate SVRM model constructed by using the methodology described in *Wang et al.* [2013]. To construct this error model, we first use the SVRM temperature model to predict the temperatures for all data points and computed $T_{\text{error}} = |T_{\text{virtual}} - T_{\text{model}}|/T_{\text{virtual}}$. We then substitute T_{virtual} with T_{error} in the training data and construct a SVRM model to predict errors.

Figure A1a suggests that MLT may not be an essential parameter to the ion temperature (note that the true MLT effect is indeed very weak as shown in Figure A1d). To investigate how excluding nonessential parameters may affect the model prediction of the dependences on other parameters, we constructed another model by using 30% of the data set but using only, r, IMF B_z, and V_{sw} as the control parameters (model 2). The prediction results from model 2 are shown in Figures A1c and A1e. Comparing with the results from the model 1, the model 2 can still predict temperature with similar correlation coefficient and NRMSD and similar variations with r, IMF B_z, and V_{sw}. On the other hand, if without knowing beforehand whether a parameter is important, then comparing the prediction results from two models, one including that parameter and the other excluding it, can indicate the importance of that parameter.

The investigation with SVRM in this study is conducted by using software LibSVM (version 3.21), a freely available and widely used implementation of SVRM, which handles all the complex technical details for this technique [*Chang and Lin*, 2011]. Detailed information for how to use LibSVM, including how to generate training and testing data files with proper formats, can be found in *Wang et al.* [2013].

Acknowledgments

We thank Vassilis Angelopoulos at University of California, Los Angeles, for helpful discussion. The work by C.-P. Wang and H.-J. Kim has been supported by NASA contract NNX14AD18G. Chao Yue gratefully acknowledges support from the NASA Living With a Star Jack Eddy Postdoctoral Fellowship Program, administered by the UCAR Visiting Scientist Programs. J.M. Weygand has been supported by NASA grants NG04GA934G and NNX15AR07G. The work by Xiangning Chu has been supported by NASA contract NNX14AN85G. The THEMIS are obtained with the Space Physics Environment Data Analysis System (SPEDAS) software (http://themis.igpp.ucla.edu/software.shtml) and are available for free. The propagated solar wind/IMF data are obtained for free on Virtual Magnetospheric Observatory (http://vmo.igpp.ucla.edu). The *AE* index from Word Data Center for Geomagnetism, Kyoto, is obtained for free on NASA CDAWeb (http://cdaweb.gsfc.nasa.gov/cdaweb/istp_public/). We acknowledge NASA contract NAS5-02099 for THEMIS and ARTEMIS, C.W. Carlson and J.P. McFadden for the use of ESA data, and K.H. Glassmeier, U. Auster, and W. Baumjohann for the use of FGM data provided under DLR contract 50-OC-0302. We thank J.H. King, N. Papatashvilli at AdnetSystems, and NASA GSFC and CDAWeb for providing the *AE* data. The LibSVRM software available for free at http://www.csie.ntu.edu.tw/cjlin/libsvm.

References

Agapitov, O., et al. (2009), Surface waves and field line resonances: A THEMIS case study, *J. Geophys. Res.*, *114*, A00C27, doi:10.1029/2008JA013553.

Angelopoulos, V., C. F. Kennel, F. V. Coroniti, R. Pellat, M. G. Kivelson, R. J. Walker, C. T. Russell, W. Baumjohann, W. C. Feldman, and J. T. Gosling (1994), Statistical characteristics of bursty bulk flow events, *J. Geophys. Res.*, *99*(A11), 21,257–21,280, doi:10.1029/94JA01263.

Auster, H. U., et al. (2008), The THEMIS fluxgate magnetometer, *Space Sci. Rev.*, *141*, 235–264, doi:10.1007/s11214-008-9365-9.

Belcher, J. W., and L. Davis Jr. (1971), Large-amplitude Alfvén waves in the interplanetary medium, 2, *J. Geophys. Res.*, *76*(16), 3534–3563, doi:10.1029/JA076i016p03534.

Birn, J., A. Runov, and M. Hesse (2014), Energetic electrons in dipolarization events: Spatial properties and anisotropy, *J. Geophys. Res. Space Physics*, *119*, 3604–3616, doi:10.1002/2013JA019738.

Borovsky, J. E., and M. H. Denton (2006), Differences between CME-driven storms and CIR-driven storms, *J. Geophys. Res.*, *111*, A07S08, doi:10.1029/2005JA011447.

Borovsky, J. E., M. F. Thomsen, and R. C. Elphic (1998), The driving of the plasma sheet by the solar wind, *J. Geophys. Res.*, *103*(A8), 17,617–17,640, doi:10.1029/97JA02986.

Bortnik, J., W. Li, R. M. Thorne, and V. Angelopoulos (2016), A unified approach to inner magnetospheric state prediction, *J. Geophys. Res. Space Physics*, *121*, 2423–2430, doi:10.1002/2015JA021733.

Chang, C. C., and C. J. Lin (2011), LIBSVM: A library for support vector machines, *ACM Trans. Intell. Syst. Technol.*, *2*, 27:1–27:27, doi:10.1145/1961189.1961199.

Dubyagin, S., N. Y. Ganushkina, I. Sillanpää, and A. Runov (2016), Solar wind-driven variations of electron plasma sheet densities and temperatures beyond geostationary orbit during storm times, *J. Geophys. Res. Space Physics*, *121*, 8343–8360, doi:10.1002/2016JA022947.

Gabrielse, C., C. Harris, V. Angelopoulos, A. Artemyev, and A. Runov (2016), The role of localized inductive electric fields in electron injections around dipolarizing flux bundles, *J. Geophys. Res. Space Physics*, *121*, 9560–9585, doi:10.1002/2016JA023061.

Gavrishchaka, V. V., and S. B. Ganguli (2001), Support vector machine as an efficient tool for high-dimensional data processing: Application to substorm forecasting, *J. Geophys. Res.*, *106*(A12), 29,911–29,914, doi:10.1029/2001JA900118.

Gkioulidou, M., C.-P. Wang, L. R. Lyons, and R. A. Wolf (2009), Formation of the Harang reversal and its dependence on plasma sheet conditions: Rice convection model simulations, *J. Geophys. Res.*, *114*, A07204, doi:10.1029/2008JA013955.

Grigorenko, E. E., E. A. Kronberg, P. W. Daly, N. Y. Ganushkina, B. Lavraud, J.-A. Sauvaud, and L. M. Zelenyi (2016), Origin of low proton-to-electron temperature ratio in the Earth's plasma sheet, *J. Geophys. Res. Space Physics*, *121*, 9985–10,004, doi:10.1002/2016JA022874.

Hsu, T.-S., and R. L. McPherron (2012), A statistical analysis of substorm associated tail activity, *Adv. Space Res.*, *50*(10), 1317–1343, doi:10.1016/j.asr.2012.06.034.

Huang, C., D.-D. Liu, and J.-S. Wang (2009), Forecast daily indices of solar activity, F10.7, using support vector regression method, *Res. Astron. Astrophys.*, *9*, 694, doi:10.1088/1674-4527/9/6/008.

Ji, E.-Y., Y.-J. Moon, J. Park, J.-Y. Lee, and D.-H. Lee (2013), Comparison of neural network and support vector machine methods for Kp forecasting, *J. Geophys. Res. Space Physics*, *118*, 5109–5117, doi:10.1002/jgra.50500.

Jiang, F., M. G. Kivelson, R. J. Walker, K. K. Khurana, V. Angelopoulos, and T. Hsu (2011), A statistical study of the inner edge of the electron plasma sheet and the net convection potential as a function of geomagnetic activity, *J. Geophys. Res.*, *116*, A06215, doi:10.1029/2010JA016179.

Kessel, R. L. (2008), Solar wind excitation of Pc5 fluctuations in the magnetosphere and on the ground, *J. Geophys. Res.*, *113*, A04202, doi:10.1029/2007JA012255.

Kim, H.-J., L. R. Lyons, S. Zou, A. Boudouridis, D.-Y. Lee, C. Heinselman, and M. McCready (2009), Evidence that solar wind fluctuations substantially affect the strength of dayside ionospheric convection, *J. Geophys. Res.*, *114*, A11305, doi:10.1029/2009JA014280.

Kim, H.-J., L. R. Lyons, J. M. Ruohoniemi, N. A. Frissell, and J. B. Baker (2012), Principal component analysis of polar cap convection, *Geophys. Res. Lett.*, *39*, L11105, doi:10.1029/2012GL052083.

Kim, H.-. J., L. Lyons, A. Boudouridis, V. Pilipenko, A. J. Ridley, and J. M. Weygand (2011), Statistical study of the effect of ULF fluctuations in the IMF on the cross polar cap potential drop for northward IMF, *J. Geophys. Res.*, *116*, A10311, doi:10.1029/2011JA016931.

Korth, H., M. Thomsen, J. Borovsky, and D. McComas (1999), Plasma sheet access to geosynchronous orbit, *J. Geophys. Res.*, *104*(A11), 25,047–25,061, doi:10.1029/1999JA900292.

Liang, H., M. Ashour-Abdalla, R. Richard, D. Schriver, M. El-Alaoui, and R. J. Walker (2014), Contrasting electron acceleration processes during two substorms, *J. Geophys. Res. Space Physics*, *119*, 5382–5400, doi:10.1002/2013JA019721.

Lin, D., C. Wang, W. Li, B. Tang, X. Guo, and Z. Peng (2014), Properties of Kelvin-Helmholtz waves at the magnetopause under northward interplanetary magnetic field: Statistical study, *J. Geophys. Res. Space Physics*, *119*, 7485–7494, doi:10.1002/2014JA020379.

Liu, D. D., C. Huang, J. Y. Lu, and J. S. Wang (2011), The hourly average solar wind velocity prediction based on support vector regression method, *Mon. Not. R. Astron. Soc.*, *413*, 2877–2882, doi:10.1111/j.1365-2966.2011.18359.x.

Liu, S., M. W. Chen, L. R. Lyons, H. Korth, J. M. Albert, J. L. Roeder, P. C. Anderson, and M. F. Thomsen (2003), Contribution of convective transport to stormtime ring current electron injection, *J. Geophys. Res.*, *108*(A10), 1372, doi:10.1029/2003JA010004.

Lyons, L. R., et al. (2009), Evidence that solar wind fluctuations substantially affect global convection and substorm occurrence, *J. Geophys. Res.*, *114*, A11306, doi:10.1029/2009JA014281.

McFadden, J. P., et al. (2008), The THEMIS ESA plasma instrument and in-flight calibration, *Space Sci. Rev.*, *141*, 277–302, doi:10.1007/s11214-008-9440-2.

Nagata, D., S. Machida, S. Ohtani, Y. Saito, and T. Mukai (2007), Solar wind control of plasma number density in the near-Earth plasma sheet, *J. Geophys. Res.*, *112*, A09204, doi:10.1029/2007JA012284.

Nykyri, K. (2013), Impact of MHD shock physics on magnetosheath asymmetry and Kelvin-Helmholtz instability, *J. Geophys. Res. Space Physics*, *118*, 5068–5081, doi:10.1002/jgra.50499.

Øieroset, M., T. D. Phan, M. Fujimoto, L. Chan, R. P. Lin, and R. Skoug (2003), Spatial and temporal variations of the cold dense plasma sheet: Evidence for a low-latitude boundary layer source? in *Earth's Low- Latitude Boundary Layer*, *Geophys. Monogr. Ser.*, vol. 133, edited by P. T. Newell and T. Onsager, pp. 53–62, AGU, Washington, D. C.

Potapov, A. S. (2013), ULF wave activity in high-speed streams of the solar wind: Impact on the magnetosphere, *J. Geophys. Res. Space Physics*, *118*, 6465–6477, doi:10.1002/2013JA019119.

Rae, I. J., et al. (2005), Evolution and characteristics of global Pc5 ULF waves during a high solar wind speed interval, *J. Geophys. Res.*, *110*, A12211, doi:10.1029/2005JA011007.

Scholkopf, B., A. Smola, R. C. Williamson, and P. L. Bartlett (2000), New support vector algorithms, *Neural Comput.*, *12*, 1207–1245, doi:10.1162/089976600300015565.

Smola, A. J., and B. Schölkopf (2004), A tutorial on support vector regression, *Stat. Comput.*, *14*, 199–222, doi:10.1023/B:STCO.0000035301.49549.88.

Sun, S., P. Ban, C. Chen, Z. Ding, and Z. Xu (2011), Low-latitude storm time ionospheric predictions using support vector machines, *Adv. Space Res.*, *47*(12), 2194, doi:10.1016/j.asr.2010.10.014.

Taylor, M. G. G. T., et al. (2012), Spatial distribution of rolled-up Kelvin-Helmholtz vortices at Earth's dayside and flank magnetopause, *Ann. Geophys.*, *30*, 1025–1035, doi:10.5194/angeo-30-1025-2012.

Terasawa, T., et al. (1997), Solar wind control of density and temperature in the near-Earth plasma sheet: WIND/GEOTAIL collaboration, *Geophys. Res. Lett.*, *24*, 935–938, doi:10.1029/96GL04018.

Tsurutani, B. T., T. Gould, B. E. Goldstein, W. D. Gonzalez, and M. Sugiura (1990), Interplanetary Alfvén waves and auroral (substorm) activity: IMP 8, *J. Geophys. Res.*, *95*(A3), 2241–2252, doi:10.1029/JA095iA03p02241.

Tsyganenko, N. A., and T. Mukai (2003), Tail plasma sheet models derived from Geotail particle data, *J. Geophys. Res.*, *108*(A3), 1136, doi:10.1029/2002JA009707.

Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, Springer-Verlag, New York.

Vapnik, V., and A. Lerner (1963), Pattern recognition using generalized portrait method, *Autom. Rem. Contr.*, *24*, 774–780.

Wang, C.-P., L. R. Lyons, T. Nagai, J. M. Weygand, and R. W. McEntire (2007), Sources, transport, and distributions of plasma sheet ions and electrons and dependences on interplanetary parameters under northward interplanetary magnetic field, *J. Geophys. Res.*, *112*, A10224, doi:10.1029/2007JA012522.

Wang, C.-P., L. R. Lyons, V. Angelopoulos, D. Larson, J. P. McFadden, S. Frey, H.-U. Auster, and W. Magnes (2008), THEMIS observations of penetration of the plasma sheet into the ring current region during a magnetic storm, *Geophys. Res. Lett.*, *35*, L17S14, doi:10.1029/2008GL033375.

Wang, C.-P., L. R. Lyons, R. A. Wolf, T. Nagai, J. M. Weygand, and A. T. Y. Lui (2009), Plasma sheet $PV^{5/3}$ and nV and associated plasma and energy transport for different convection strengths and AE levels, *J. Geophys. Res.*, *114*, A00D02, doi:10.1029/2008JA013849.

Wang, C.-. P., L. R. Lyons, T. Nagai, J. M. Weygand, and A. T. Y. Lui (2010), Evolution of plasma sheet particle content under different interplanetary magnetic field conditions, *J. Geophys. Res.*, *115*, A06210, doi:10.1029/2009JA015028.

Wang, C.-. P., M. Gkioulidou, L. R. Lyons, R. A. Wolf, V. Angelopoulos, T. Nagai, J. M. Weygand, and A. T. Y. Lui (2011), Spatial distributions of ions and electrons from the plasma sheet to the inner magnetosphere: Comparisons between THEMIS-Geotail statistical results and the Rice convection model, *J. Geophys. Res.*, *116*, A11216, doi:10.1029/2011JA016809.

Wang, Y., D. G. Sibeck, J. Merka, S. A. Boardsen, H. Karimabadi, T. B. Sipes, J. Šafránková, K. Jelínek, and R. Lin (2013), A new three-dimensional magnetopause model with a support vector regression machine and a large database of multiple spacecraft observations, *J. Geophys. Res. Space Physics*, *118*, 2173–2184, doi:10.1002/jgra.50226.

Weimer, D. R. (2004), Correction to "Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance tech- nique," *J. Geophys. Res.*, *109*, A12104, doi:10.1029/2004JA010691.

Weimer, D. R., D. M. Ober, N. C. Maynard, M. R. Collier, D. J. McComas, N. F. Ness, C. W. Smith, and J. Watermann (2003), Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance technique, *J. Geophys. Res.*, *108*(A1), 1026, doi:10.1029/2002JA009405.

Wing, S., J. R. Johnson, P. T. Newell, and C.-. I. Meng (2005), Dawn-dusk asymmetries, ion spectra, and sources in the northward interplanetary magnetic field plasma sheet, *J. Geophys. Res.*, *110*, A08205, doi:10.1029/2005JA011086.

Yang, J., F. R. Toffoletto, and R. A. Wolf (2016), Comparison study of ring current simulations with and without bubble injections, *J. Geophys. Res. Space Physics*, *121*, 374–379, doi:10.1002/2015JA021901.

Yue, C., C.-P. Wang, L. Lyons, Y. Wang, T.-S. Hsu, M. Henderson, V. Angelopoulos, A. T. Y. Lui, and T. Nagai (2015a), A 2-D empirical plasma sheet pressure model for substorm growth phase using the Support Vector Regression Machine, *J. Geophys. Res. Space Physics*, *120*, 1957–1973, doi:10.1002/2014JA020787.

Yue, C., C.-P. Wang, Y. Nishimura, K. R. Murphy, X. Xing, L. Lyons, M. Henderson, V. Angelopoulos, A. T. Y. Lui, and T. Nagai (2015b), Empirical modeling of 3-D force-balanced plasma and magnetic field structures during substorm growth phase, *J. Geophys. Res. Space Physics*, *120*, 6496–6513, doi:10.1002/2015JA021226.