A new multivariate time series data analysis technique: Automated detection of flux transfer events using Cluster data

H. Karimabadi,[1,2] T. B. Sipes,[1] Y. Wang,[3] B. Lavraud,[4,5] and A. Roberts[3]

Received 2 March 2009; accepted 14 April 2009; published 23 June 2009.

[1] A new data mining technique called MineTool-TS is introduced which captures the time-lapse information in multivariate time series data through extraction of global features and metafeatures. This technique is developed into a JAVA-based data mining software which automates all the steps in the model building to make it more accessible to nonexperts. As its first application in space sciences, MineTool-TS is used to develop a model for automated detection of flux transfer events (FTEs) at Earth's magnetopause in the Cluster spacecraft time series data. The model classifies a given time series into one of three categories of non-FTE, magnetosheath FTE, or magnetospheric FTE. One important feature of MineTool-TS is the ability to explore the importance of each variable or combination of variables as indicators of FTEs. FTEs have traditionally been identified on the basis of their magnetic field signatures, but here we find that some plasma variables can also be effective indicators of FTEs. For example, the perpendicular ion temperature yields a model accuracy of ~93%, while a model based solely on the normal magnetic field B_N yields an accuracy of ~95%. This opens up the possibility of searching for more unusual FTEs that may, for example, have no clear B_N signature and create a more comprehensive and less biased list of FTEs for statistical studies. We also find that models using GSM coordinates yield comparable accuracy to those using boundary normal coordinates. This is useful since there are regions where magnetopause models are not accurate. Another surprising result is the finding that the algorithm can largely detect FTEs, and even distinguish between magnetosheath and magnetospheric FTEs, solely on the basis of models built from single parameters, something that experts may not do so straightforwardly on the basis of short time series intervals. The most accurate models use a combination of plasma and magnetic field variables and achieve a very high accuracy of prediction of ~99%. We explain the high detection accuracies both in terms of the existence of clear physical signatures of FTEs (for the majority of cases) and in terms of the capability of the data mining technique to explore the data set in a much more thorough fashion than expert human eyes. A list of 1222 FTEs from Cluster data during years 2001–2003 is provided as auxiliary material.

Citation: Karimabadi, H., T. B. Sipes, Y. Wang, B. Lavraud, and A. Roberts (2009), A new multivariate time series data analysis technique: Automated detection of flux transfer events using Cluster data, *J. Geophys. Res.*, *114*, A06216, doi:10.1029/2009JA014202.

1. Introduction

[2] Many data analysis problems in space sciences stand to benefit from the application of data mining techniques. One such class of problems is automated detection of events. Examples of events include boundary (e.g., bow shock, current sheet) crossings and flux transfer events (FTEs), among others. The traditional approach for search algorithms in space sciences has been based on rote methods where one assumes very specific signatures and/or threshold conditions and writes code to find matching conditions in the data. For example, *Khurana and Schwarzl* [2005] used the reversals in the radial component of the magnetic field to identify current sheet crossings. *Kawano and Russell* [1996] used isolated bipolar signatures satisfying a predetermined set of conditions including its duration and the peak-to-peak amplitude of the bipolar field component B_N to search for FTEs in ISEE 1 data. Such schemes are most effective in cases where the events of interest have simple and well established signatures. Rote methods lack computational "intelligence" as they are bound by the specific criteria hardwired by the user. In contrast, data mining algorithms autonomously determine criteria and conditions to uniquely identify an event.

[1]SciberQuest, Inc., Del Mar, California, USA.
[2]Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, California, USA.
[3]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[4]Centre d'Etude Spatiale des Rayonnements, Université de Toulouse, UPS, Toulouse, France.
[5]UMR5187, Centre National de la Recherche Scientifique, Toulouse, France.

Copyright 2009 by the American Geophysical Union.
0148-0227/09/2009JA014202

[3] Here we introduce a new multivariate time series analysis technique called MineTool-TS and as its first application, we apply it to the automated detection of flux transfer events in the Cluster data. Besides offering high accuracy of the resulting predictive models, a key advantage of MineTool-like approach is that it makes data mining more accessible to nonexperts, by offering a self-contained step by step procedure for model building.

[4] Flux transfer events (FTEs) are thought to be the result of patchy, intermittent reconnection at the Earth's magnetopause, i.e., the boundary that separates the solar wind and shocked magnetosheath plasma from the magnetospheric cavity [*Russell and Elphic*, 1979; *Berchem and Russell*, 1984]. Although several models have been invoked to account for their formation [e.g., *Lee and Fu*, 1985; *Southwood et al.*, 1986; *Sonnerup*, 1987; *Scholer*, 1988; *Pu et al.*, 1990; *Raeder*, 2006; *Karimabadi et al.*, 2005, 2007a], the presence of a bipolar signature in the normal component of the magnetic field B_N (relative to the "locally planar" magnetopause boundary) is the common signature that is typically used for their identification. However, there is also evidence of FTEs without the characteristic bipolar perturbation [e.g., *Le et al.*, 2008]. This indicates the need to be able to find FTEs on the basis of other signatures than just the bipolar signature in B_N.

[5] Here we use MineTool-TS and a labeled list of FTEs to develop a search algorithm for FTEs. For a given time series, the algorithm produces one of three outputs, non-FTE, magnetosheath FTE, and magnetospheric FTE. The distinction between the two "types" of FTEs is based on which side of the magnetopause the spacecraft encounters the FTE and does not refer to a different physical process. The observed signatures do show differences between the two types of FTEs [e.g., *Elphic*, 1995; *Le et al.*, 1999] and hence we will keep the distinction of the two types here. We emphasize that the three output detection is a much more challenging problem in data mining than the two output case.

[6] Our goal was threefold. (1) Develop an automated scheme for detection of FTEs. (2) Experiment with different combinations of variables, including models based on only one variable (e.g., V_y or B_L, etc.) to determine the relative role of each variable for identification of FTEs. This study addresses such questions as the following: Can FTEs be detected in the data on the basis of their plasma signature alone, can other components of magnetic field besides B_N be good indicators of FTEs, and can FTEs be detected without using boundary normal coordinates? (3) Run the algorithm on an existing FTE list to check for consistency and accuracy of the list.

[7] The results show many surprises regarding the influence of different magnetic and plasma variables on the detection of FTEs. For example, we found that the algorithm is able to distinguish between magnetospheric and magnetosheath FTEs solely on the basis of the magnetic field data. The paper is structured as follows. Section 2 describes the MineTool-TS algorithm. For readers more interested in how to use the algorithm, they can refer to Appendix A where it is shown that by virtue of automating the modeling steps, no knowledge of the algorithms is required to use this tool. In fact using the default settings, running the models is as simple as opening a file and clicking a button. The workings of MineTool-TS, behind the scenes and masked from the user, are discussed in Appendix B. Section 3 describes the application of MineTool-TS to detection of FTEs in the Cluster data. Section 4 compares the model predictions with current understanding of FTEs. Summary and conclusion are presented in section 5.

2. MineTool-TS Algorithm

2.1. Background

[8] In a static data, each data point is an independent observation. An example is the position of magnetopause as a function of solar wind conditions. In contrast, in a time series the neighboring points define a pattern and one is often interested in an automated fashion to detect such patterns in the data.

[9] Time series analysis has become one of the most important branches of mathematical statistics and a variety of techniques have been developed. The techniques range from a single time series forecasting such as the ARIMA method [*Mills*, 1990] to multivariate time series analysis. The latter, which is the focus of the present work, is used when one is interested in understanding and modeling the interactions among a group of time series variables. Just like in any other classification problem, we are given examples of labeled data in order to build a predictive model. Historically, Hidden Markov Models (HMMs), recurrent Artificial Neural Networks (recurrent ANNs) and Dynamic Time Warping (DTW) have been used to build predictive models of multivariate time series data for classification tasks [*Rabiner and Juang*, 1986; *Meyers and Rabiner*, 1981; *Rumelhart and McClelland*, 1986]. Even though these techniques are useful for certain tasks, they have several disadvantages which make them impractical for large data sets. In case of HMMs, for example, the number of parameters that needs to be set and examined is very large. HHMs also make several major assumptions not readily available in a real-world scientific data set. Recurrent ANNs suffer from several of the same problems as HMMs and require the user to experiment and choose many parameters and decide on the appropriate network architecture.

2.2. Our Approach

[10] Recently we developed a new data mining algorithm for *static* data called MineTool [*Karimabadi et al.*, 2007] that automates the modeling process to a simple click of a button upon selecting a data file. The algorithm is, however, flexible and does offer the user the capability to modify the default parameters to experiment with model accuracy. MineTool could be used for time series data if one could replace the time series by a static data consisting of variables that capture its essential and interesting features (e.g., number of zero crossings, slope, extreme values). One simple and obvious idea for reduction of time series data is to replace the time series by its statistical measures such as the mean, standard deviation, minimum and maximum values. We have found this approach to be unsatisfactory and yielding poor accuracy (not shown). Instead we use a more sophisticated approach to extract features from multivariate time series data that yields much higher accuracy. *Kadous* [2002] used metafeatures in extracting information

from multivariate time series data, for classification purposes. In this technique, called TClass, various simple features of time series data are extracted, and then segmented using directed segmentation or k-means clustering. This produces the "synthetic events" data set which is then fed into a decision tree method [*Quinlan*, 1993]. Kadous achieved fairly to very accurate classification results in the domains of ECG (electrocardiograph) diagnosis (72% accuracy) and the sign language recognition (98% accuracy). We build on TClass for the physics domain application by (1) improving the feature collection and segmentation paradigms, (2) adding more features, and (3) utilizing MineTool [*Karimabadi et al.*, 2007] as the classification method of choice, which has the capacity to outperform standard data mining tools such as decision trees, artificial neural networks [*Ripley*, 1996] and support vector machines [*Cortes and Vapnik*, 1995].

[11] We call the extension of MineTool to time series data MineTool-TS. The main change is in the inclusion of an extra step that "flattens" the time series data into static data. By this, we mean that the time lapse information stored in multiple rows is converted in a single row containing all the important time varying information. Thus, all the advantages of MineTool, including ease of use, automation of modeling process, high accuracy, and analytical models, are preserved. Readers more interested in learning how to use MineTool-TS rather than the algorithmic issues can skip the remaining parts of this section and should refer to Appendix A. Here we describe our procedure for treatment of time series data:

[12] 1. Metafeature Collection: These are features in the data that capture the local changes in a time series. Currently MineTool-TS uses four metafeatures: (1) increasing, (2) decreasing, (3) plateau, and (4) bipolar. These are discussed in Appendix A. Others metafeatures can be easily added.

[13] 2. Global Feature Collection: As the name indicates, global features describe the global properties of the time series. Examples include global minimum and maximum, mean, and number of zero crossings.

[14] 3. Clustering of the Collected Features: The collection of new metafeature and global feature variables would potentially increase the number of input variables quite a bit, as for each of the input variables a total of four metafeature and four global feature variables is collected to represent the time varying information. There are several significant risks to having too many inputs. One is that potentially useful candidates get overlooked simply because there are too many variables to evaluate. Another is that if a systematic routine for evaluating and including variables is used, it can lead to overfit [*Karimabadi et al.*, 2007]. Finally, many candidate variables are likely to be redundant, which can cause difficulties for the estimation routines. MineTool-TS deals with these issues by clustering the similar novel feature variables into groups or clusters. In this manner, all the similar features are replaced by one average (i.e., the cluster center or "centroid") feature. This decreases the number of variables and increases the predictive power of the feature variables that convert the time series data into a static set. Unlike the previous efforts that either use simple directed segmentation or k-means clustering [*Kadous*, 2002] which requires a prior specification of the "k" number of clusters, in MineTool-TS we apply the Expectation Maximization (EM) clustering [*Hartley*, 1958; *Dempster et al.*, 1977; *McLachlan and Krishnan*, 1997] that offers an optimal number of clusters. In this way, the metafeature variables are grouped into an optimal number of clusters, which assures that no information is lost in reducing the number of variables (such as in predetermining that k = 3 where possibly the optimal number of clusters is 5 as given by the EM algorithm). In this way MineTool-TS performs the no-loss-of-information clustering of the novel feature variables, for the purposes of having an increased predictive power and reduced number of input variables into the intermediate (static) data set.

[15] 4. Intermediate Data Set Creation: This step is used to ensure that all the information collected using the feature extraction method is converted into a static data set. This so-called intermediate data set contains the important time-varying information from each of the univariate time series examples. Again, we improve on the previous work in this area by recording not just presence or absence of a feature cluster in a univariate time series; rather we create an entry that effectively describes the specific metafeature. The global feature entries are then added to the cluster centroids (i.e., major value group's average).

[16] By using the above time series preprocessing, we have created a static intermediate data set that still contains all the important time-varying information. Standard MineTool can then be applied to this data set. The procedure for using MineTool-TS, somewhat similar to a user manual, is provided in Appendix A. In the following, we focus on the specific details regarding the selection and preparation of Cluster data as well as the discussion of the resulting models.

3. Application to Cluster Data

3.1. Preparation of Cluster Data

[17] There are a number of procedures to prepare the data before it can be used by the data mining algorithm.

3.1.1. Data Selection

[18] The Cluster mission was launched in 2000. It consists of four identical satellites with an interspacecraft separation varying over the course of the mission between 100 and several thousand km. It has an elliptic polar orbit with a perigee of 4 R_E, an apogee of 19.6 R_E and a period of ~58 h. The plane of the orbit precesses clockwise looking down from the north, and it extends outside of the magnetopause from near the end of one year to around summer of the next year, during which FTEs can be observed. Cluster magnetic field observations from the Fluxgate Magnetometer (FGM) [*Balogh et al.*, 1997] and ion observations from the Cluster Ion Spectrometry (CIS) instrument [*Rème et al.*, 2001] at the spin resolution (4 s) were used for magnetopause crossing and FTE identifications.

[19] A stretched *Shue et al.* [1998] magnetopause model [*Wang et al.*, 2005] is used for the magnetic field LMN boundary normal coordinate transformations. This leads to a total of 14 input variables: B_x, B_y, B_z, B_L, B_M, B_N, $|B|$, N_p, V_x, V_y, V_z, $|V|$,$T_{\|}$, and $T_{\perp}$. B_{xyz} are magnetic field components in Geocentric Solar Magnetic (GSM) coordinates; B_{LMN} are those in the LMN coordinates. V_{xyz} are the components of the ion velocity in GSM coordinates. N_p is

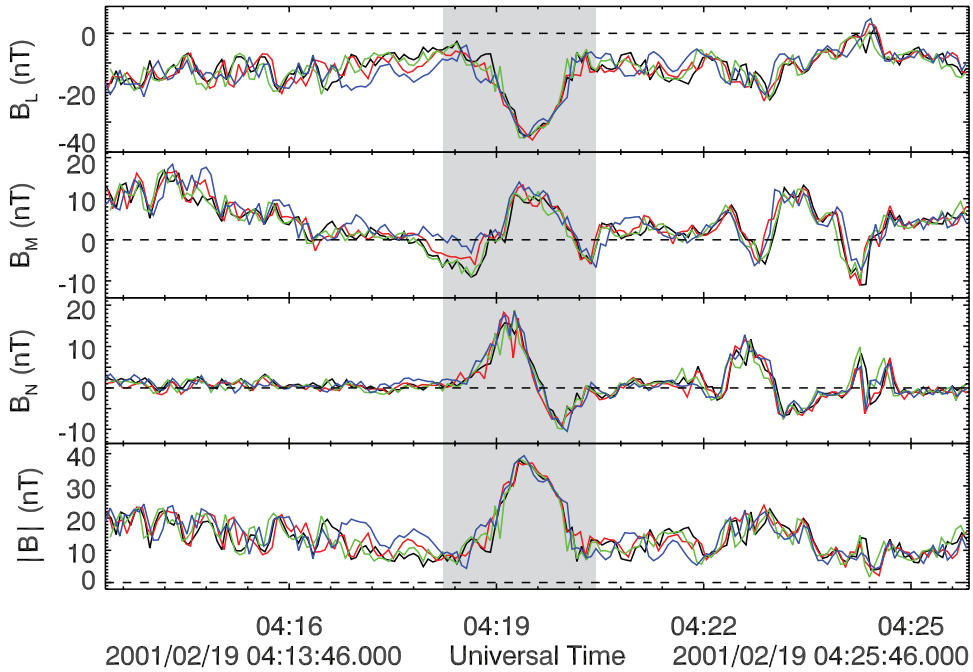

Figure 1. An example of a magnetospheric FTE observed at ~0420 UT at (5.8, 0.8, 10.2) R_E in GSM coordinates. Different colors refer to different Cluster satellites. Note the clear B_N bipolar signature and the total magnetic field enhancement.

the total ion density, and $T_{\|}$ and $T_{\perp}$ are the parallel and perpendicular (to the magnetic field) ion temperatures. We also included the additional variables T_p (the total ion temperature) where $T_p = (T_{\|} + 2T_{\perp})/3$ and $T_{\perp}/T_{\|}$ in order to determine their importance as FTE indicators. Note that MineTool-TS enables us to experiment with different combinations of these variables and their effects on the model accuracy.

3.1.2. Time Series Generation

[20] It is important that all time series have the same temporal span and that all the variables included as input have the same temporal resolution so that they line up in the time series. In our case, both the magnetic field and plasma data have the same resolution. Otherwise, additional data treatment would be required before data mining can be performed (e.g., interpolation).

[21] A total of 615 time series, corresponding to intervals containing magnetosheath, magnetospheric, or non-FTEs, are generated with each time series consisting of 14 input variables and spanning 6 min with 89 different time-lapse measurements. The choice of 6 min is rather arbitrary aside from the fact that it has to be long enough to accommodate the typical duration of an FTE, which is typically lower than 1 min [cf. *Wang et al.*, 2006]. The data is then "cleaned" to remove any gaps or anomalies (see *Press et al.* [2007] for a general overview). By anomaly, we mean values that are beyond the expected range for a given variable. Such anomalies are found through a simple search algorithm.

3.1.3. Labeling Data

[22] For each time series, we add an integer entry between 0 and 2 that indicates whether it is non-FTE (0), magnetosheathFTE (1) or magnetosphereFTE (2). It is important to note that one does not need to specify the start and finish of the FTE or even whether there is more than one FTE in a given time series for the labeled data set. For example, while rare in a 6 min interval, we did have cases with three FTEs within the same time series. This makes the creation of the labeled data set much easier and it is apparently sufficient given the high accuracy of the resulting models. The output of the model is an integer between 0 and 2 for each time series. But the algorithm does not provide the start and finish of the FTE nor does it specify how many FTEs are within the 6 min time interval. Rather it specifies whether there are any FTEs within a time series and if so, what type. We are currently working on another technique to pinpoint the start and finish of the FTE in a time series and will report the results elsewhere.

[23] The expert labeling of the data set used for training the algorithms was performed through visual inspection of data based on the following considerations. An event was labeled as FTE if it exhibits a clear bipolar signature of B_N and an associated enhancement in $|B|$. It was also required that an FTE clearly isolates itself from its surroundings, which avoids identifying FTEs in highly oscillating field structures. Whether an FTE is a magnetosheath or magnetosphere FTE was determined by two methods that yielded similar results. In the first method, we checked whether Cluster is moving into or out of the magnetosphere and whether the FTE is observed before or after the Cluster magnetopause crossing. For example, if Cluster is moving into the magnetosphere and the FTE is observed before Cluster magnetopause crossing, the FTE is a magnetosheath FTE. The second method was based on the consideration of the absolute values of the surrounding plasma density and temperature that are respectively substantially lower and higher in the case of magnetospheric FTEs. Both methods were used jointly for the final magnetosphere/magnetosheath FTE labels used in this study. For more details about LMN

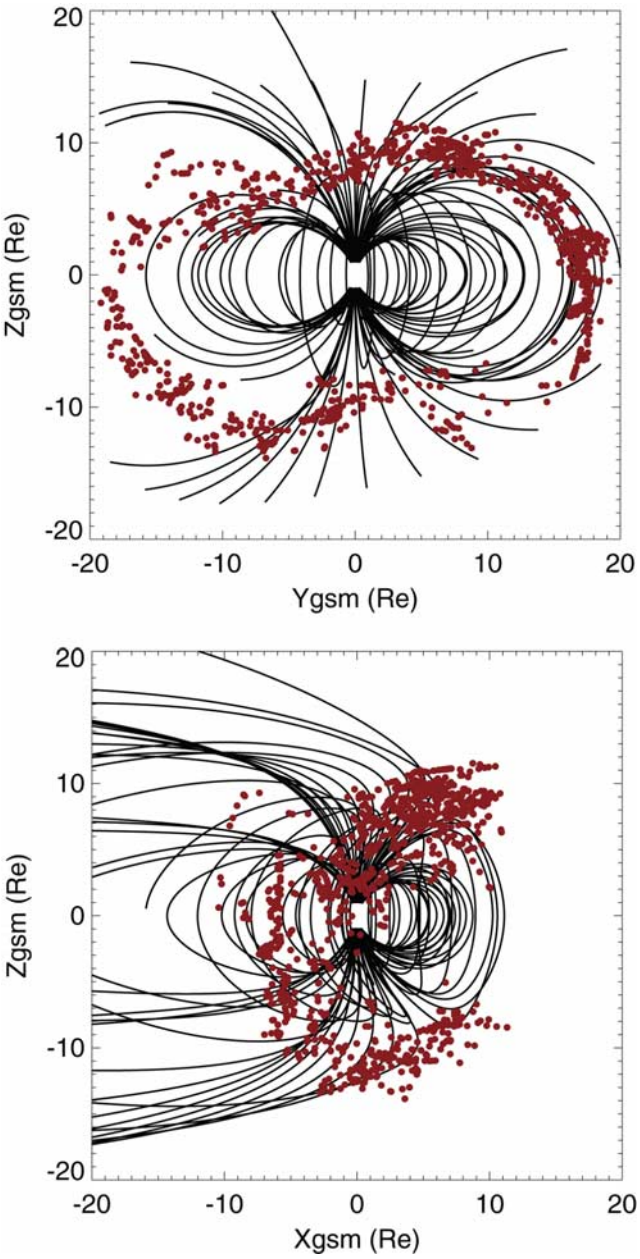

Figure 2. The spatial distribution of the Cluster FTEs used in this study.

transformation and FTE identification, please refer to *Wang et al.* [2005]. An example of a magnetospheric FTE is shown in Figure 1 (shaded region). In this case the FTE has a clear bipolar signature and an associated enhancement in the total magnetic field. There is also a less clear FTE at ~0423 UT [*Wang et al.*, 2005].

[24] Note that aside from the label, none of the information about the movement of Cluster or the temporal sequence of FTE observation relative to the magnetopause crossing is captured in the data set used for training the algorithm. But as we will show shortly, the algorithm can still make the distinction between the two types of FTEs with very high accuracy.

[25] Our data set consisted of 195 non-FTEs, 267 magnetosheath FTEs, and 153 magnetospheric FTEs. The non-FTEs were selected from segments where spacecraft did not see FTEs.

[26] A given FTE may be observed by only one or up to all four Cluster spacecraft but the data from each spacecraft is treated as an independent time series. In the case of labeled FTEs, we center the time series on the FTE observations time with 3 min before and 3 min after it. Figure 2 shows the spatial distribution of our FTE list which includes FTEs from both near-equatorial crossings as well as high-latitude crossings. Note that $B_y \sim B_N$ for low-latitude flank FTEs, $B_z \sim B_N$ for high-latitude FTEs and $B_x \sim B_N$ for near-equatorial subsolar FTEs. Owing to the orbit of Cluster, our list does not include any proper near-equatorial subsolar FTEs.

[27] MineTool automatically breaks the data into a training set and a test or hold out set. The training set consisted of two thirds of the examples, and the hold out set of one third of the instances. This choice of split is a standard evaluation method in data mining. Only the training data is visible during the training and modeling phases, and the test set is held out to evaluate the models.

3.2. Choice of Metafeatures and Global Features

[28] Time series information of each input variable was converted into a static (i.e., flattened) meta data through the use of three meta features including increases, decreases, and plateau (i.e., typical time derivates) as well as three global features consisting of mean, global maximum and global minimum. Other metafeatures such as bipolar signature can be easily incorporated. We excluded the bipolar signature metafeature for two reasons. First, the bipolar signature of B_N requires transformation to the boundary normal coordinates whereas here one of the issues that we wanted to consider was whether FTE detection would be possible in standard GSM coordinates. Second, our intent was to have the algorithm identify specific features and relations between the input variables that uniquely classify the event into one of the three categories rather than "hardwire" our own prescription of the FTE identifier.

[29] In order to explore the sensitivity of model accuracy to particular choice of variables, we built fifteen single-variable and thirteen multivariable models.

3.3. Calculation of Error Estimates

[30] MineTool incorporates several ways to quantify the model errors. Evaluation measures play an important role in data mining and are used not only to compare the accuracy of the different models but also as goals to optimize in constructing learning models. There are a variety of measures that have been proposed [e.g., *Caruana and Niculescu-Mizil*, 2004] such as the correlation coefficient (CC), the root mean squared error (RMSE), the mean absolute error (MAE), and the false and true positive rates. One common way to visualize the results is the confusion matrix where each column of the matrix represents the number of events in a predicted class, while each row represents the number of events in an actual class. The diagonal entries would then represent the correct predictions while the off-diagonal terms represent the cases for which the algorithm confuses the type. However, it is impractical to display the confusion matrix for all the models that we are considering here. Thus, we distill the information in the confusion matrix into three

Table 1. Comparison of Accuracy of FTE Detection Models Based on Individual Variables in Order to Determine the Relative Importance of Physical Variables as FTE Indicators

Three-Tier Classification		Classification of FTE Versus Non-FTE					Classification of Msheath Versus Msphere FTE		
Model	Correctly Classified (P_3)	Correctly Classified (FTE Versus Non-FTE) (P_1)	Number of False Non-FTEs		Number of False FTEs		Correctly Classified (Msth Versus Msph FTE) (P_2)	Number of False Msth FTEs	Number of False Msph FTEs
			msth	msph	msth	msph			
$T_\perp/T_\parallel$	59%	83%	7	0	23	4	67%	50	0
N_p	77%	92%	2	0	13	1	79%	30	1
B_{tot}	91%	95%	0	0	8	2	94%	0	9
B_z	92%	93%	9	0	14	0	98%	0	3
T_p	92%	97%	0	0	6	0	93%	0	10
B_L	92%	96%	0	0	7	1	95%	3	5
$T_\parallel$	93%	97%	0	0	5	1	94%	0	9
$T_\perp$	93%	99%	0	0	2	0	91%	11	2
V_x	93%	98%	0	0	4	0	93%	7	4
B_x	93%	94%	0	0	11	1	99%	0	2
V_z	94%	97%	0	0	6	0	95%	0	7
B_y	94%	98%	0	0	4	1	95%	0	8
B_N	95%	99%	0	0	2	0	94%	3	6
V_y	96%	98%	0	0	3	2	98%	0	3
B_M	96%	99%	0	0	2	1	97%	0	5

quantities P_1, P_2, and P_3. P_1 is the percentage of cases that the algorithm correctly distinguishes between non-FTEs and FTEs, regardless of their type, P_2 is the percentage of correctly classified type of FTEs among the cases that were correctly classified as FTE by the algorithm, and P_3 is an overall performance measure determined by the percentage of correct classification of all three event types. To calculate the accuracy percentages, we use the following notation:

- tN_{tot} total number of events in the test data set;
- N_{FTE} total number of FTEs in the test data set;
- $F_{Non\text{-}FTE}$ otal number of FTEs which were misclassified as non-FTEs;
- F_{FTE} total number of non-FTEs which were misclassified as FTEs;
- $F_{Msheath}$ total number of magnetospheric FTEs that were misclassified as magnetosheath FTEs.
- $F_{Msphere}$ total number of magnetosheath FTEs that were misclassified as magnetospheric FTEs.

[31] The P's are calculated as

$$P_1 = \frac{N_{tot} - F_{Non-FTE} - F_{FTE}}{N_{tot}}, \tag{1a}$$

$$P_2 = \frac{N_{FTE} - F_{Msheath} - F_{Msphere}}{N_{FTE}}, \tag{1b}$$

$$P_3 = \frac{N_{tot} - F_{Non-FTE} - F_{FTE} - F_{Msheath} - F_{Msphere}}{N_{tot}}. \tag{1c}$$

To express these in percentage, we multiply them by hundred (e.g., see Table 1). Since the FTE label has three distinct values: 0 (non-FTE), 1 (magnetosheath FTE) and 2 (magnetospheric FTE), we use the standardized cutoff values to calculate the percent of correctly classified instances. To calculate the percentage of correctly classified instances, all the instances with the model prediction value of less than 0.5 are set to 0, with the value larger than 1.5 are set to 2, and the rest are set to 1. Then, the instances that match the actual (label) values are labeled as correctly classified, the ones that differ are marked as misclassified, and the percentages are calculated accordingly.

3.4. Test Accuracy of Labeled Set

[32] Once the model is built from the training set of data (2/3 of the data), we run the model on the test (1/3 of the data) labeled data set and compare the algorithm's label with the expert label for each event. We then compile a list of cases where the algorithm label did not agree with the expert label. The mismatch list usually consists of cases where the algorithm got the label wrong as well as cases where the expert label was wrong. In our case, we found 18 cases that upon further examination turned out to be due to error in the expert label. In the process we found several new FTEs in the data that had been tagged as non-FTEs by experts. An example is shown in Figure 3 where the initial expert label had it classified as a non-FTE but the algorithm tagged it as a magnetosheath FTE. Upon further examination we concluded that the algorithm is correct and there is indeed an FTE between 2002/12/08 0609:49.659 and 0610:38.085. Recall that the expert label of whether an FTE is a magnetosheath or magnetosphere FTE was determined by checking whether Cluster is moving into or out of the magnetosphere and whether the FTE is observed before or after the Cluster magnetopause crossing. An equivalent method was based on the consideration of the absolute values of the surrounding plasma density and temperature that are respectively substantially lower and higher in the case of magnetospheric FTEs. The algorithm has no knowledge of the movement or positioning of the Cluster spacecraft, but presumably in models that do include N_p and T_p the algorithm could key on their absolute values as a way to distinguish between the types of FTEs. But as we will show shortly, even without N_p and T_p, the algorithm does manage to distinguish between the two types of FTEs with high accuracy, which is somewhat surprising.

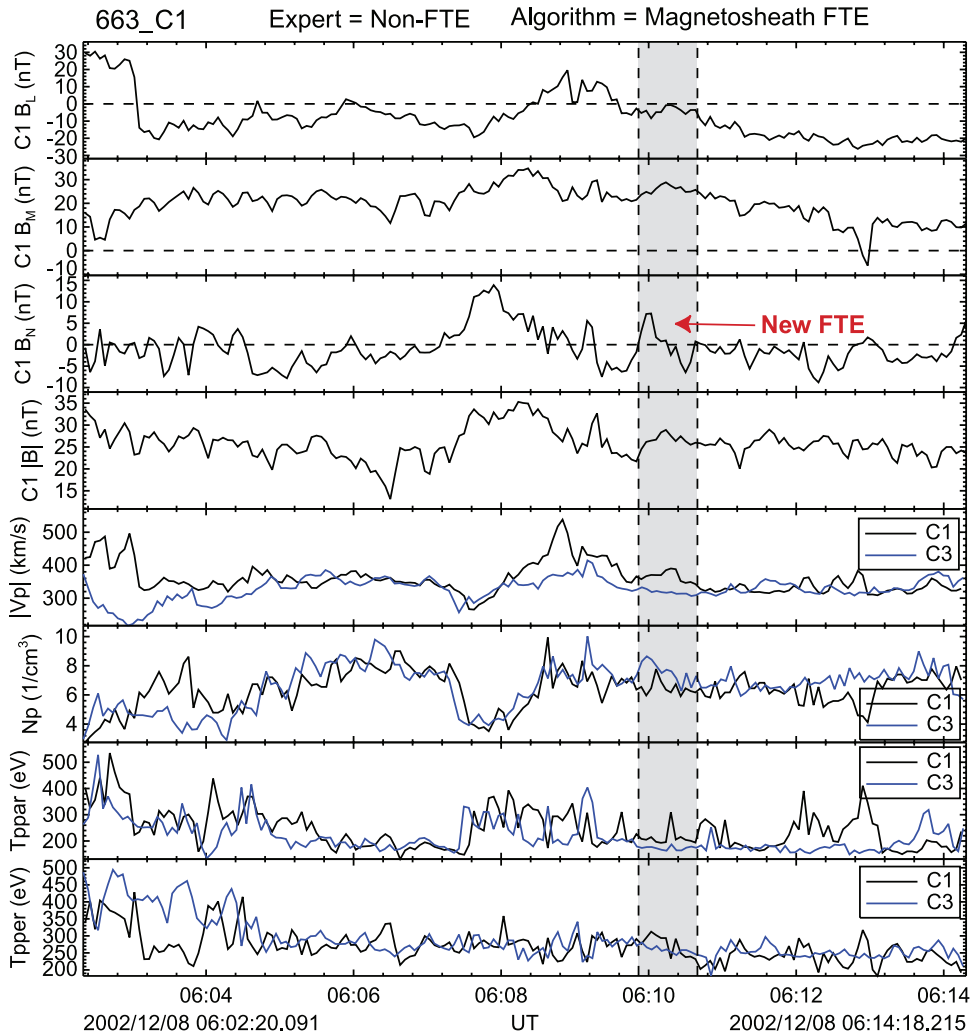

Figure 3. A segment of Cluster observations showing a structure that the algorithm identified as a magnetosheath FTE, whereas the original labeled set had it classified as a non-FTE. Upon further examination the algorithm was found to have the correct classification.

3.5. Resulting Models

[33] Flux transfer events have been identified primarily on the basis of their bipolar signature in the normal component of the magnetic field, B_N. Since different processes can produce a bipolar signature, usually other variables such as the total magnetic field are also used in identifying an FTE. To distinguish between magnetospheric and magnetosheath FTE, additional information has been required such as the ion density and temperature, whether Cluster is moving into or out of the magnetosphere, and whether the FTE is observed before or after the Cluster magnetopause crossing. Note that the information about the movement of Cluster or the relative position of magnetopause is not presented to the algorithm. Our goal in this section is to compare the model accuracy as a function of different combination of variables to assess their importance as unique FTE identifiers.

[34] In the foregoing, for each model, we varied the variable transformation parameters in MineTool-TS, which regulate the type of transformations of the input variables to be included into modeling. These include the basic transformations of the input, the logistic and hyperbolic tangent transformation of the input as well as the Artificial Neural Network-like transformations of the input [*Karimabadi et al.*, 2007].

3.6. Single Variable Models

[35] We start by models constructed using the time series from only one input variable at a time. Since B_N has been used primarily in the identification of FTEs, it is interesting to compare the model accuracy of other models against that built on the basis of B_N. Table 1 shows the comparison of single variable models listed from the top in order of increasing accuracy as measured by P_3 (equation (1c)). It is evident that aside from temperature anisotropy and density, all other input variables in Table 1 perform well as distinguishers between the three classes (non-FTE, magnetosheath FTE, and magnetospheric FTE) with accuracies in the range of ~91% to 96%. Some of the other important results of Table 1 are as follows.

3.6.1. As Indicator of FTEs (P_1)

[36] If only interested in finding FTEs in the data, and not concerned about distinguishing between magnetosheath

Table 2. Variations in Model Accuracy as a Function of Different Sampling of the Training and Test Data for Four Models[a]

Model	$P_3)_{set1}$	$P_3)_{set3}$	$P_3)_{set3}$	$P_3)_{set4}$	Variation in P_3
$T_\perp/T_\parallel$	59.0%	57.1%	59.0%	57.6%	1.9%
$\mathbf{B_{tot}}$	90.7%	92.2%	90.7%	91.2%	1.5%
$\mathbf{B_x}$	93.2%	94.6%	93.2%	94.6%	1.4%
$\mathbf{V_y}$	96.1%	95.1%	96.1%	96.6%	1.5%

[a]Set 1 is the same sample data used for the development of the models in Table 1. The other three sets consist of different sampling of the labeled data.

and magnetospheric FTEs, then all individual variables considered proved to be good indicators. The model based on temperature anisotropy has the lowest accuracy of ~83% and that based on either B_N or $T_\perp$ has the highest accuracy of 99%. It is very surprising that some of the variables, such as V_x, are very good indicators of FTEs. There is a general lack of systematic behavior of the flows at and around FTEs since they largely vary both in the magnetosheath (depending on distance from the nose) and in the magnetosphere (e.g., dayside plasma sheet, boundary layers and cusp offer a wide variety of flows). If true, then our result suggests that there are distinct flow patterns within/around the FTEs that make the identification of FTEs possible.

[37] In all models, the main source of error in P_1 was in misclassification of non-FTEs as magnetosheath FTEs. Similarly while some of the magnetosheath FTEs were misclassified as non-FTEs, none of the models mistook a magnetospheric FTE for a non-FTE. This indicates that it may be more difficult to distinguish a magnetosheath FTE in the data, at least on the basis of the use of a single variable. This likely stems from the fact that the magnetosheath is a much more variable region, in terms of plasma and field properties, than the magnetospheric region near the magnetopause.

3.6.2. As Indicator of the Type of FTE (P_2)

[38] Of the pool of events correctly identified as FTEs, most variables served as very good indicators of the types of FTEs. The one exception is the temperature anisotropy which misclassified a large number of magnetospheric FTEs as magnetosheath FTEs and did not do much better than a random selection. This may be explained by the fact that the magnetosheath, the dayside magnetosphere and the inside of FTEs all show perpendicular temperature anisotropies, thus rendering the distinction unlikely. However, the total as well as parallel and perpendicular temperatures alone appear to be good indicators of the types of FTEs. The local enhancement in the FTEs and the average values in the surrounding plasma may explain this. The temperatures are better indicators than the density, but the reason for that is unclear.

3.6.3. Overall Performance (P_3)

[39] Temperature anisotropy $T_\perp/T_\parallel$ leads to a model with poor accuracy that is not much better than random selection. Most of the error is in its misclassification of magnetospheric FTEs as magnetosheath FTE (this may be explained along the lines discussed in the previous paragraph). Density does much better than $T_\perp/T_\parallel$ as an indicator of the type of FTE but it misclassifies a large number of non-FTEs as FTEs. V_y, B_M and B_N yield the most accurate models with 95–96% correctly classified. Of the plasma variables, V_y yields the highest accuracy. Although it is expected that a tangential flow change (which would generally show up in the V_y component) occurs inside FTEs owing to local stress balance following from reconnection, it is overall unclear why this particular component does so well.

[40] It is known that some FTEs do not have signatures in the plasma bulk velocity or the total magnetic field [e.g., *Elphic*, 1995]. And more recently an FTE was reported with no clear bipolar signature in B_N. Let us consider such an example. Clearly a single-variable model based on B_N alone would have difficulty to correctly identify such an FTE. Some of the misclassified cases in the single-variable models may in fact be cases where the FTEs have no clear signature in those variables. That is why using models with combination of variables would be advantageous for finding unusual FTEs.

3.6.4. Overall Accuracy of the Models

[41] Table 1 shows that most models except those based on temperature anisotropy and density have a tight range in accuracy (P_3) from ~90%–96%. The natural question is how meaningful are the differences of a few percentages among the models and can one draw physical conclusions from them. If we define the error estimate as the degree of variation in the model accuracies as we choose different samplings of the data, then the expected error in using 66–33% split evaluation method is typically about 2%. Note that this ~2% variation is independent of the accuracy of the model with a model having a 55% model accuracy would have a similar error estimate of ~2% as that having a model accuracy of 90%. To verify this expectation, we used different segments of the data for our training and test sets in the 66–33% split evaluation method and built three additional models for each of the following single variables V_y, B_x, B_{tot}, and $T_\perp/T_\parallel$. The results are shown in Table 2, which shows the variation in P_3 with three different sampling of the data in addition to that originally used in Table 1 (marked as set 1). As expected, the variations in P_3 are less than 2% for all cases. One should, however, exercise caution in drawing general conclusions about the physics of FTEs based on variations of only several percentages among models. For instance, the present work is based on a particular labeled set. Using a different criteria for preparing the labeled set (e.g., clear bipolar signature and a well defined variations in B_M rather than an increase in B_{tot}) is expected to lead to somewhat different relative accuracies of the models.

3.7. Multivariable Models

[42] Table 3 shows the comparison of models using combination of input variables. All models yield classification scores P_3 of 93% and higher with the best model achieving ~99% accuracy which is comparable to the most accurate of the single variable models. The robustness of the results with accuracy almost independent of the particular combination of variables indicates that FTEs have strong, distinctive signatures in both their magnetic field and plasma structures. Note also that of the pool of correctly identified FTEs, all models do an amazingly good job of classifying the types of FTEs ($P_2 > 97\%$). The fact that the models can distinguish so effectively between the magnetosheath and magnetospheric FTEs is partly due to the fact that the surrounding plasma density and temperature are respectively substantially lower and higher in the case of magnetospheric FTEs. However, this cannot be the only

Table 3. Comparison of Accuracy of FTE Detection Models Using Combination of Variables

Three-Tier Classification		Classification of FTE Versus Non-FTE					Classification of Msheath Versus Msphere FTE		
Model	Correctly Classified (P_3)	Correctly Classified (FTE Versus Non-FTE) (P_1)	Number of False Non-FTEs		Number of False FTEs		Correctly Classified (Msth Versus Msph FTE) (P_2)	Number of False Msth FTEs	Number of False Msph FTEs
			msth	msph	msth	msph			
B_x,B_{tot}	93%	95%	0	0	10	0	97%	0	4
B_x,B_y,B_z, B_{tot}	95%	95%	0	0	10	0	100%	0	0
B_x,B_y,B_z, B_{tot},T_p,N_p	96%	96%	0	0	8	0	99%	0	1
B_N,B_{tot}	96%	98%	0	0	5	0	98%	0	3
B_N,B_{tot}	96%	98%	0	0	5	0	98%	0	3
B_N,B_{tot},N_p	96%	97%	0	0	6	0	99%	0	2
B_x,B_y,B_z, B_{tot},V_x,V_y, V_z, T_p,N_p	96%	96%	0	0	8	0	100%	0	0
$T_\perp$,$T_\parallel$	97%	99%	1	0	1	0	97%	0	5
B_L,B_M,B_N,B_{tot},T_p,N_p	97%	99%	0	0	3	0	97%	0	4
B_L,B_M,B_N, B_{tot}	98%	98%	0	0	4	0	99%	0	1
B_x,B_y,B_z, B_{tot},V_x,V_y, V_z, N_p, $T_\perp$,$T_\parallel$	98%	99%	0	0	3	0	99%	0	2
B_L,B_M,B_N, B_{tot}, N_p, V_x,V_y,V_z,T_p	98%	98%	0	0	4	0	100%	0	0
B_L,B_M,B_N, B_{tot},N_p	99%	99%	0	0	3	0	100%	0	0
B_x,B_{tot}	93%	95%	0	0	10	0	97%	0	4
B_x,B_y,B_z, B_{tot}	95%	95%	0	0	10	0	100%	0	0

reason since the models including just the magnetic field such as the B_x, B_{tot} also yield very high accuracy. One possible explanation is that magnetosheath and magnetospheric FTEs may have systematic differences in their structures.

[43] Another interesting finding in Table 3 is that using the three components of the magnetic field in either GSM coordinates or boundary normal coordinates yields comparable model accuracies of ~95% and ~98%, respectively. This is important since there are regions in the magnetosphere (e.g., high-latitude dayside magnetopause) where the magnetopause models are not as accurate and the transformation to boundary normal coordinates becomes problematic.

4. Physics of Model Results

[44] Overall, the model accuracies are amazingly high. We may not provide simple, unambiguous explanations for all of the results reported in Tables 1 and 3. In this section we only discuss in detail some of the results which we think are of particular interest.

[45] We noted that the total temperature alone gives high detection accuracy (92%). Although at first surprising, it appears in fact obvious that the model can notice a localized temperature variation, characteristic of the core of FTEs, and at the same time can make the difference between the magnetosphere and magnetosheath on the basis of the average value in the surrounding. Also, it is not necessarily surprising that the temperature alone does almost as good as the magnetic field components alone, for instance. Indeed, large fluctuations in magnetic fields unrelated to FTEs are likely more frequent than large fluctuations in temperatures, in the magnetosheath in particular.

[46] Another interesting and rather surprising result is that B_L and B_M alone give comparable accuracies to B_N alone. Previous studies of FTEs had revealed no universal trends in B_L and B_M. *Paschmann et al.* [1982] showed considerable variation of B_L and B_M from event to event. *Rijnbeek et al.* [1984] looked for a correlation between the sign of B_M and the occurrence of FTEs. They found that the crossings with southward (negative B_L) fields tend to be usually associated with a duskward (negative B_M) orientation, with no corresponding bias for the crossings with positive or intermediate B_L fields. However, *Kawano and Russell* [1996] showed that there were only 933 out of 1246 FTEs showing rotational polarity in the MN plane. In contrast to these works, the algorithm is able to decipher trends in the data that are not easily discernible to the human eye.

[47] Furthermore, all magnetic field components alone have high accuracies in distinguishing between magnetospheric and magnetosheath FTEs, i.e., even without information from the surrounding plasma properties (density or temperature in particular). This might be explained by the fact that the magnetosphere has a strong and steady northward component, while the magnetosheath is typically more variable and generally southward during FTEs. Remnants of the field orientation and variability are likely present even in individual magnetic field components (i.e., in their background value), and which the algorithm might be able to pick while human eyes might not.

[48] The algorithm's ability to outperform human eye in uncovering hidden relations and dependencies in the data is partly due to the fact that the human eye is ideally suited for examination of a still frame. With an increase in the size of data (either in terms of number of variables in a still frame and/or the presence of many frames), the data mining techniques quickly gain advantage. In case of time series classification it is nearly impossible for the human eye to "see" the differences between the time series and collect it into a coherent pattern. The algorithm does this by taking into account all the increases/decreases over all the examples of one category (i.e., FTEs versus non-FTEs), then generalizes what was collected (by using clustering), and then only account for the product of gradient and average value of the event.

5. Summary and Conclusion

[49] We described a new time series analysis technique called MineTool-TS which automates the modeling process. We then used it to develop automated models for detection of flux transfer events (FTEs) at Earth's magnetopause

using Cluster data. The data including both FTEs and non-FTEs used for developing the models are provided as auxiliary material Data Sets S1 and S2.[1] Also included is the list of 1222 FTEs (Data Set S3).

[50] The best model produced about 99% accuracy with the only cases of misclassification being 3 cases of non-FTEs which were misclassified as magnetosheath FTEs. The resulting model can be used in three distinct manners. First, it can be run on search for new FTEs in previously unexplored data. Second, by running different models and comparing their accuracy one can determine the relative role of each input variable to unique identification of FTEs. Third, one can run the model on any previously compiled list of FTEs to check for consistency and accuracy of the list.

[51] When applied to a time series, the model produces one of three outputs, non-FTE, magnetosheath FTE, and magnetospheric FTE. A number of surprises and interesting results came out of this study:

[52] 1. The surprisingly good accuracies of FTE detection from many single and combined magnetic and plasma parameters show much more intrinsic and distinct features of FTEs than previously thought. Such features are extremely difficult, if not impossible, for human experts to decipher but are quite apparent using MineTool-TS. This finding makes it possible to detect FTEs beyond human expert capabilities.

[53] 2. The models based on GSM variables had comparable accuracies to those based on boundary normal coordinates. Since magnetopause model is not as accurate in locations such as high latitude, the use of GSM coordinates which does not require any assumption of a magnetopause model is advantageous.

[54] 3. The model is able to distinguish between magnetosheath and magnetospheric FTEs even without knowledge of the relative motion of spacecraft with respect to the magnetopause, and for a number of models even without the surrounding density and temperature information. It suggests that the average field orientations in the magnetosphere and magnetosheath constitute enough information to make the difference, and that is true even for only one magnetic field component.

[55] 4. Although previous studies of FTEs had found no universal trends in B_L and B_M, the single variable models of B_L and B_M yielded very accurate results, suggesting that there may indeed be distinctive patterns in these magnetic components that may not be readily discernable in visual inspection of time series.

[56] 5. The velocity information does not seem to be necessary information in the identification of FTEs considered here. Its inclusion did not yield noticeable improvement in the accuracy of the models with multiple parameters. It is noticeable, though, that some velocity components (e.g., V_Y) can do a good job in identifying FTEs on their own.

[57] 6. To the best of our knowledge all previous detections of FTEs have had to rely on the pattern of at least two variables. The bipolar signature in B_N is typically used as a necessary condition along with at least one more variable such as B_{tot} or B_M. This is the first time it has been shown that single variables can lead to high-accuracy detection of FTEs.

[58] 7. The model results are overall unexpectedly high for almost all parameters, and demonstrate the power of MineTool-TS algorithm in analyzing time series data in magnetospheric physics. We should mention that our application of MineTool in other areas such as genomic is also yielding equally accurate results.

[59] These results serve as validation of MineTool-TS as an efficient and accurate modeling tool for automated detection of events such as FTEs in magnetospheric physics. One obvious bias has clearly been the requirement that FTEs exhibit *clear*, bipolar signature in B_N. Using this as the necessary condition, *Elphic* [1995] proposed a taxonomy of FTEs based on different behavior of plasma and magnetic signatures among the different classes. A significant finding here is that one can find FTEs using other indicators besides B_N. This enables us to search for FTEs that may not have a clear or any variation in B_N or may lack variations in other expected parameters such as B_{tot}. In fact, more recently an example of an FTE with no clear bipolar signature in B_N was reported [*Le et al.*, 2008]. Thus, an important extension of our work is in creation of a more comprehensive and unbiased list of FTEs for statistical studies. Another potentially interesting application of our work is to investigate how the technique may similarly help in flux rope identification in the solar wind and in the magnetotail.

Appendix A

[60] We are in the process of developing a wiki page under sciberquest.com website where users can download the program and check for updates. In the mean time, the source code can be requested from the first author. Here we describe how to use MineTool-TS. The user is presented with three options upon launching MineTool-TS: (1) create a model, (2) run a previously created model, and (3) open workspace. Use of workspace enables the user to save their work and come back to it later to finish or make modification to a completed run. Figure A1 shows the flowchart for creating a new model. The first step, currently outside of MineTool-TS, is to prepare the data in a form suitable for data mining. This involves "cleaning" the data (e.g., removing gaps, interpolating, etc.) [e.g., *Press et al.*, 2007] and labeling it. As we explained earlier, in the case of FTE classification problem considered here, we labeled each time series by adding an extra entry to the time series indicating whether it was a non-FTE, magnetosheath FTE, or a magnetospheric FTE. Next release of MineTool-TS will include some data manipulation routines for interpolation and removing of out of bound values for convenience. Once the data is ready to be read, the user can launch MineTool and select the button to create a new model (not shown). MineTool-TS has default settings that are designed to provide accurate results in most cases. Thus, the creation of a model can be as simple as reading in the data and pushing the "create model" button. Upon completion of the calculations by the algorithm, it displays model summary and if the results are acceptable (e.g., error estimates are small), the model can be saved. If not, the user can change the parameters until a satisfactory result is obtained. This is

[1]Auxiliary materials are available at ftp://ftp.agu.org/apend/ja/2009ja014202.

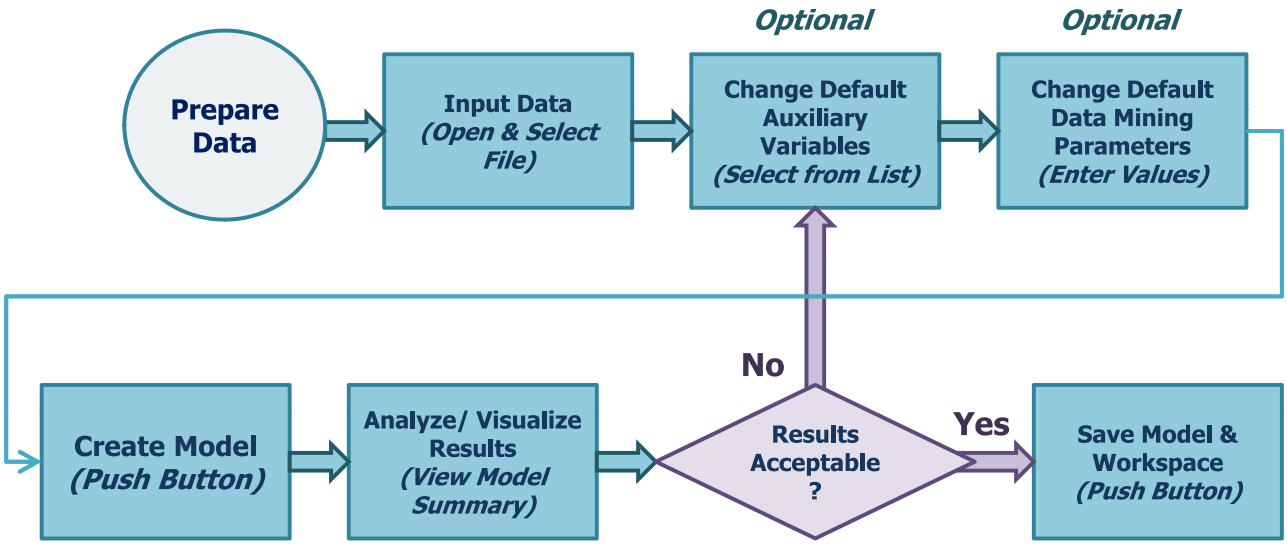

Figure A1. The steps in data mining in MineTool-TS from a user's perspective. Using the default parameters, the user can create a model by reading in a preformatted data file and clicking a button to run. MineTool-TS also offers the user the ability to change the default settings to experiment with model accuracy.

done by modifying the settings in the two boxes labeled as "optional" in Figure A1.

A1. List of Auxiliary Variables

[61] In addition to the original variables, additional variable transformations can be included. MineTool-TS currently includes the following options with others easily added: (1) level 2 transformations ($x^{\alpha}y^{\beta}$, $\alpha,\beta = -1,0,1$), (2) level 3 transformations ($x^{\alpha}y^{\beta}z^{\gamma}$, $\alpha,\beta,\gamma = -1,0,1$), (3) level 4 transformations ($x^{\alpha}y^{\beta}z^{\gamma}w^{\delta}$, $\alpha,\beta,\gamma,\delta = -1,0,1$), (4) exponential, log, tanh, cosine, and sine transformations, and (5) ANN-like transformations.

[62] For a detailed discussion of these parameters, the reader is referred to *Karimabadi et al.* [2007]. In addition to these input parameters, there are time series–specific parameters that can be selected:

[63] 1. Increasing Metafeature: An increasing metafeature is recorded for all the consecutive rising time series measurements. MineTool-TS allows for a small amount of noise to be ignored, so that the true increasing events are captured. For each increasing event, we record its start point, duration, gradient and average value, so that the increasing events can be used for analysis, comparison and data mining later in the algorithm.

[64] 2. Decreasing Metafeature: For each decreasing event, similarly to the increasing events, we record starting point, duration, gradient (which is negative in this case) and average value.

[65] 3. Plateau Metafeature: A plateau metafeature is recorded for all the consecutive non-changing time series measurements. For each plateau, we record its start point, duration and the (average) value.

Table B1. MineTool-TS Shielding of the User From the Complexity of the Algorithms[a]

MineTool GUI Tab and Action	User Input	Action of the Algorithm
Input Data	Select the input file. Select the variables to include.	Read in the input file. Check for correct format. Read in the input variable values.
Change Default Auxiliary Variables	Choose from a list of transformations.	Creation of basic input transformations. Creation of advanced transformations, if selected. Ordering of all the transformations.
	If time series data: Choose from a list of meta and global features.	If time series data: Collection of selected metafeatures. Collection of selected global features. Creation of the feature data set. Clustering of similar variables. Creation of the static, feature data set that captures time varying info.
Change Default Data Mining Parameters	Change the colinearity parameter.	Ordering of all the input transformation based on the correlation to the output.
Create Model	Press the Create Model button.	Creation of training and test sets. Selection of variables and auxiliary transformations for inclusion in the model, for each value of λ. Evaluation of all intermediate models, for one value of λ. Selection of the final model. Calculation of all evaluation scores for the final model.
Analyze Results	Check measures of model accuracy.	Display the final model and evaluation scores in textual and graphical formats.
Save Model & Workspace	Push the save button.	Saves both the workspace as well as analytical model.

[a]Each user input in Figure A1 activates a series of operations behind the scenes.

[66] 4. Bipolar Signature Metafeature: A bipolar signature metafeature is recorded for all the consecutive time series measurements that increase, decrease and cross the zero, and increase again. This is a complex metafeature that consists of several simple metafeatures.

[67] 5. Global Minimum: For each single variable, the global minimum feature extracts the minimum value of all of the observations of that variable and records it as the global minimum feature for that input channel.

[68] 6. Global Maximum: The maximum value of all the sequential time observations is recorded as the global maximum feature for that variable.

[69] 7. Mean: The average value of all the sequential time observations is collected as the global mean feature for that specific variable.

[70] 8. Number of Zero Crossings: Last, the number of zero crossings occurring during the recorded measurements is written down as the number of zero crossings global feature.

A2. Data Mining Parameters

[71] In addition to the transformations, the accuracy of MineTool-TS can be affected by the choice of colinearity (redundancy) parameter λ. This parameter controls which variable transformations are included into the candidate model. See *Karimabadi et al.* [2007] for a more description of this parameter. λ can range from 0.1 to 1 with the default value set to 0.5. MineTool-TS currently uses the 66–33% split for creation of training and test sets but we plan to include other methods such as cross-validation in the near future. This would be another factor that could alter the model accuracy.

Appendix B

[72] The ease of use of MineTool-TS is masked by complex computations carried out by the algorithm behind the scenes. Table B1 lists the specific actions of the algorithms for each of the user inputs in Figure A1.

[73] **Acknowledgments.** The authors are grateful to the Cluster Ion Spectrometry (CIS) and FluxGate Magnetometer (FGM) instrument teams for the use of the data sets. The research of H. Karimabadi and T. Sipes was supported by a NASA LWS grant NNH06CD22C and a NASA SBIR Program contract. The authors acknowledge useful discussions with David Sibeck and Jan Merka.

[74] Wolfgang Baumjohann thanks Michael A. Hapgood and another reviewer for their assistance in evaluating this paper.

References

Balogh, A., et al. (1997), The Cluster magnetic field investigation, *Space Sci. Rev.*, *79*, 65–91, doi:10.1023/A:1004970907748.

Berchem, J., and C. T. Russell (1984), Flux-transfer events on the magnetopause: Spatial distribution and controlling factors, *J. Geophys. Res.*, *89*, 6689–6703, doi:10.1029/JA089iA08p06689.

Caruana, R., and A. Niculescu-Mizil (2004), Data mining in metric space: An empirical analysis of supervised learning performance criteria, in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 69–78, Assoc. for Comput. Mach., New York.

Cortes, C., and V. Vapnik (1995), Support-vector networks, *Mach. Learning*, *20*, 273–297, doi:10.1023/A:1022627411411.

Dempster, A., N. Laird, and D. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc., Ser. B*, *39*(1), 1–38.

Elphic, R. C. (1995), Observations of flux transfer events: A review, in *Physics of the Magnetopause*, *Geophys. Monogr. Ser.*, vol. 90, edited by P. Song, B. U. O. Sonnerup, and M. F. Thomsen, pp. 225–233, AGU, Washington, D. C.

Hartley, H. (1958), Maximum likelihood estimation from incomplete data, *Biometrics*, *14*, 174–194, doi:10.2307/2527783.

Kadous, M. W. (2002), Temporal classification: Extending the classification paradigm to multivariate time series, Ph.D. thesis, School of Comput. Sci. and Eng., Univ. of New South Wales, Sydney, New South Wales, Australia.

Karimabadi, H., W. Daughton, and K. B. Quest (2005), Antiparallel versus component merging at the magnetopause: Current bifurcation and intermittent reconnection, *J. Geophys. Res.*, *110*, A03213, doi:10.1029/2004JA010750.

Karimabadi, H., T. B. Sipes, H. White, M. Marinucci, A. Dmitriev, J. K. Chao, J. Driscoll, and N. Balac (2007), Data mining in space physics: MineTool algorithm, *J. Geophys. Res.*, *112*, A11215, doi:10.1029/2006JA012136.

Kawano, H., and C. T. Russell (1996), Survey of flux transfer events observed with the ISEE 1 spacecraft: Rotational polarity and the source region, *J. Geophys. Res.*, *101*, 27,299–27,308, doi:10.1029/96JA02703.

Khurana, K. K., and H. K. Schwarzl (2005), Global structure of Jupiter's magnetospheric current sheet, *J. Geophys. Res.*, *110*, A07227, doi:10.1029/2004JA010757.

Le, G., J. T. Gosling, C. T. Russell, R. C. Elphic, M. F. Thomsen, and J. A. Newbury (1999), The magnetic and plasma structure of flux transfer events, *J. Geophys. Res.*, *104*(A1), 233–245, doi:10.1029/1998JA900023.

Le, G., et al. (2008), Flux transfer events simultaneously observed by Polar and Cluster: Flux rope in the subsolar region and flux tube addition to the polar cusp, *J. Geophys. Res.*, *113*, A01205, doi:10.1029/2007JA012377.

Lee, L. C., and Z. F. Fu (1985), A theory of magnetic flux transfer at the Earth's magnetopause, *Geophys. Res. Lett.*, *12*, 105–108, doi:10.1029/GL012i002p00105.

McLachlan, G., and T. Krishnan (1997), *The EM Algorithm and Extensions*, John Wiley, New York.

Meyers, C. S., and L. R. Rabiner (1981), A comparative study of several dynamic time-warping algorithms for connected word recognition, *Bell Syst. Tech. J.*, *60*(7), 1389–1409.

Mills, T. C. (1990), *Time Series Techniques for Economists*, Cambridge Univ. Press, Cambridge, U. K.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2007), *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Cambridge Univ. Press, Cambridge, U. K.

Paschmann, G., G. Haerendal, I. Papamastoraki, N. Sckopke, S. J. Bame, J. T. Gosling, and C. T. Russell (1982), Plasma and magnetic field characteristics of magnetic flux transfer events, *J. Geophys. Res.*, *87*, 2159–2168, doi:10.1029/JA087iA04p02159.

Pu, Z. Y., P. T. Hou, and Z. X. Liu (1990), Vortex induced tearing mode instability as a source of flux transfer events, *J. Geophys. Res.*, *95*, 18,861–18,869, doi:10.1029/JA095iA11p18861.

Quinlan, J. R. (1993), *C4.5: Programs for Machine Learning*, Morgan Kaufmann, New York.

Rabiner, L. R., and B. H. Juang (1986), An introduction to hidden Markov models, *IEEE Mag. Acoust. Speech Signal Process.*, *3*(1), 4–16.

Raeder, J. (2006), Flux transfer events: 1. Generation mechanism for strong southward IMF, *Ann. Geophys.*, *24*, 381–392.

Rème, H., et al. (2001), First multispacecraft ion measurements in and near the Earth's magnetosphere with the identical Cluster ion spectrometry (CIS) experiment, *Ann. Geophys.*, *19*, 1303–1354.

Rijnbeek, R. P., S. W. H. Cowley, D. J. Southwood, and C. T. Russell (1984), A survey of dayside flux transfer events observed by ISEE 1 and ISEE 2 magnetometers, *J. Geophys. Res.*, *89*, 786–800, doi:10.1029/JA089iA02p00786.

Ripley, B. D. (1996), *Pattern Recognition and Neural Networks*, Cambridge Univ. Press, Cambridge, U. K.

Rumelhart, D. E., and J. L. McClelland (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, *Foundations*, MIT Press, Cambridge, Mass.

Russell, C. T., and R. C. Elphic (1979), ISEE observations of flux transfer events at the dayside magnetopause, *Geophys. Res. Lett.*, *6*, 33–36, doi:10.1029/GL006i001p00033.

Scholer, M. (1988), Magnetic flux transfer at the magnetopause based on single X line bursty reconnection, *Geophys. Res. Lett.*, *15*, 291–294, doi:10.1029/GL015i004p00291.

Shue, J.-H. (1998), Magnetopause location under extreme solar wind conditions, *J. Geophys. Res.*, *103*, 17,691–17,700, doi:10.1029/98JA01103.

Sonnerup, B. U. O. (1987), On the stress balance in flux transfer events, *J. Geophys. Res.*, *92*, 8613–8620, doi:10.1029/JA092iA08p08613.

Southwood, D. J., M. A. Saunders, M. W. Dunlop, W. Mierjedrzejowicz, and R. P. Rijnbeek (1986), A survey of flux transfer events recorded by the UKS spacecraft magnetometer, *Planet. Space Sci.*, *34*, 1349–1359, doi:10.1016/0032-0633(86)90071-1.

Wang, Y. L., et al. (2005), Initial results of high-latitude magnetopause and low-latitude flank flux transfer events from 3 years of Cluster observations, *J. Geophys. Res.*, *110*, A11221, doi:10.1029/2005JA011150.

Wang, Y. L., R. C. Elphic, B. Lavraud, M. G. G. T. Taylor, J. Birn, C. T. Russell, J. Raeder, H. Kawano, and X. X. Zhang (2006), Dependence of flux transfer events on solar wind conditions from 3 years of Cluster observations, *J. Geophys. Res.*, *111*, A04224, doi:10.1029/2005JA011342.

H. Karimabadi, Department of Electrical and Computer Engineering, University of California, San Diego, 9500 Gilman Drive, MC 0407, La Jolla, CA 92093, USA. (homak@sciberquest.com)

B. Lavraud, Centre d'Etude Spatiale des Rayonnements, Université de Toulouse, UPS, 9 Avenue du Colonel Roche, F-31028 Toulouse CEDEX 4, France.

A. Roberts and Y. Wang, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

T. B. Sipes, SciberQuest, Inc., 2130 Via Mar Valle, Del Mar, CA, USA.