J. Astron. Space Sci. 20(2), 101–108 (2003)

RECURRENT PATTERNS IN DST TIME SERIES

Hee-Jeong Kim[1†], Dae-Young Lee[1], and Wongyu Choe[2]

[1] Department of Astronomy and Space Science,
College of Natural Sciences and Institute for Basic Science Research,
Chungbuk National University, Chungju, Korea

[2] AIT Inc., Seocho-dong 1564-11, Seoul, 137-874, Korea

email: heekim@chungbuk.ac.kr, dylee@chungbuk.ac.kr, wgchoe@aitcorp.co.kr

(Received April 26, 2003; Accepted May 30, 2003)

ABSTRACT

This study reports one approach for the classification of magnetic storms into recurrent patterns. A storm event is defined as a local minimum of Dst index. The analysis of Dst index for the period of year 1957 through year 2000 has demonstrated that a large portion of the storm events can be classified into a set of recurrent patterns. In our approach, the classification is performed by seeking a categorization that minimizes thermodynamic free energy which is defined as the sum of classification errors and entropy. The error is calculated as the squared sum of the value differences between events. The classification depends on the noise parameter T that represents the strength of the intrinsic error in the observation and classification process. The classification results would be applicable in space weather forecasting.

Keywords: Dst index, recurrent patterns, storm classification

1. INTRODUCTION

When the coupling of the solar wind to the earth magnetosphere becomes intense, a magnetic storm develops which can be characterized by Dst time series (Burton et al. 1975, Su & Konradi 1975). Dst (storm-time disturbance) index is defined as the instantaneous worldwide average of the equatorial magnetic field (H-component) disturbance. Larger negative value of Dst index indicates a more intense magnetic storm.

The typical time evolution pattern of solar wind parameters are very complicated, sometimes described as turbulent and chaotic. Therefore, at a first sight, the time evolution of Dst index would be equally complicated and unpredictable. In the examination of the history of Dst values, however, we can easily find numerous couples of the Dst patterns that look quite similar. As an example, two geomagnetic storms observed on Oct. 11, 1980 and Sep. 26, 1981 are plotted together in Figure 1. The similarity of the Dst time series around the storm minimum (marked at t=0) is evident. Another such pair of events are plotted in Figure 2, where the two events are recorded on around Jul. 15, 1959 and Nov. 9, 1991, respectively. These observations give a rise to a simple question about Dst variations : Are there any recurrent patterns in Dst time series?

Finding the recurrent patterns are closely related to the categorization of storm events. If we find a set of events which can be characterized by a common recurrent Dst pattern, we can classify these

[†]corresponding author

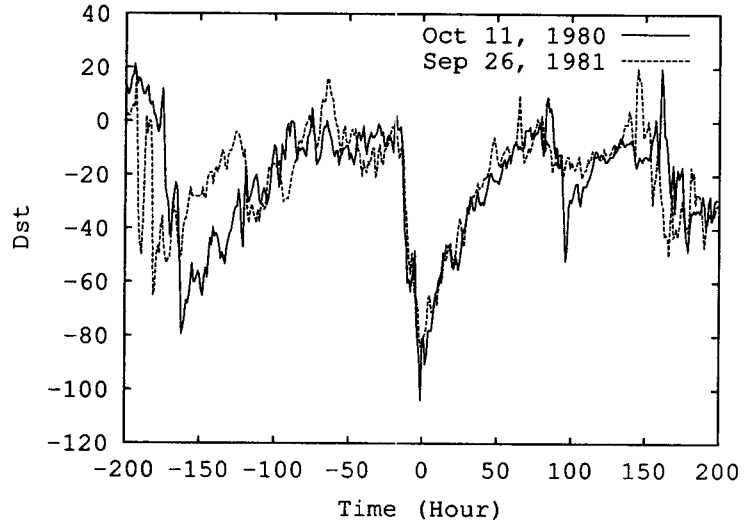

Figure 1. Two Dst events exhibiting close time evolution, which can be classified as belonging to a same recurrent pattern.

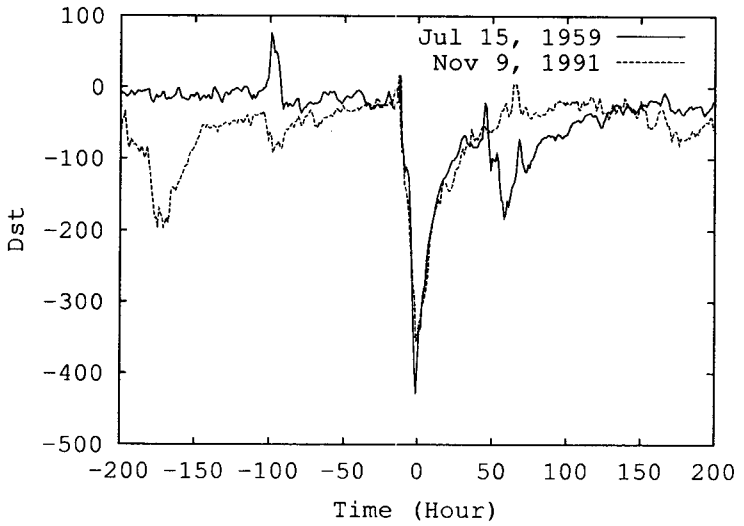

Figure 2. Another pair of Dst events which can be classified into a pattern different from that in Figure 1.

events into a category. Repeating the classification, we may end up with a (partial) categorization of the historically observed Dst events. Then a future event can be expected to belong to one of these categories. Examination of the category-wide properties rather than studying individual events could lead us to a more robust measure of the magnetospheric dynamics.

The categorization can be further utilized in space weather forecasting. By examining input solar wind parameters for the events belonging to a certain category, one can pick out category-specific measure and event-specific measure of the input parameters. For instance, we may be able to find a characteristic function of input parameters which yields different values for events belonging to different categories, but gives out (roughly) the same value for events belonging to the same category. Therefore, if we knew the value of characteristic function for given input solar wind, we could predict which category the forthcoming event will belong to.

This paper proposes one statistical scheme for the classification of Dst patterns based on statistical analysis. Following sections will discuss how the classification scheme works and which factors must be taken into account for the meaningful classification of storm events.

2. PROCEDURE

We collected Dst data from WDC-II database of Kyoto university for the period of years 1957 $\sim$ 2000. An event is defined as a local minimum of Dst index, whose value is less than -50 nT. Total 807 such events are identified during the period.

In order to categorize the events, we first need to define an error function that measures the difference between two events. A simple value-squared error is introduced as:

$$Error(E_i, E_j) = \sum_{-10<t'<50} \left(Dst(t_i + t') - Dst(t_j + t')\right)^2 \tag{1}$$

In other words, $Error(E_i, E_j)$ is given as the sum of squared value difference between the i-th event (E_i) and j-th event (E_j) over an interval spanning from 10 hours before the Dst minima to 50 hours after the minima. t_i and t_j denotes the time at the Dst minimum for events E_i and E_j, respectively. There can be many different choices for the error function, as will be discussed later. However, in this work, we focus on describing our classification process, rather than discussing validity of a specific choice of an error function. And a simple classification scheme requires only the definition of the error function.

Next is to set an error threshold ϵ. Any two events with the error less than ϵ are considered to belong to a same category. ϵ can be an any reasonably small number, or may be chosen from the distribution of all pairwise errors between events, so that the probability of observing error less than ϵ is very small, say, 0.001% . By picking an event E_i and repeating pairwise error computation, we end up with a set of events that belong to the same category with E_i. Repeating this process for each unclassified events, we can obtain a complete categorization of the Dst events. Note that many events may remain unclassified (or classified into a pattern of its own, i.e., a "trivial" pattern). The number of unclassified patterns would increase for smaller error threshold ϵ.

When we find the events belonging to a category j, we can define a template pattern P of the category as the average of all the events as follows:

$$P_j(t) = \frac{1}{n_j} \sum_{1<k<n_j} Dst(t_k + t) \tag{2}$$

where n_j is the number of the events in the category j and t_k is the k-th event in that category.

With the template pattern $P_j(t)$, one can define the intra-class error U_j, i.e., the classification error inside a pattern class. Using the error function defined in equation (1), the intra-class error U_j is given as

$$U_j = \sum_{1 \le k \le n_j} Error(P_j, E_k) \tag{3}$$

The global error associated with a classification scheme is then written as

$$U = \sum_{1 \le j \le c} U_j \tag{4}$$

where c is the number of categories that are found. So, the classification problem is reduced to an optimization problem, finding a partitioning of the event set which minimizes the global error U. First, assume that every event is of its own pattern, so that $c = N$ and $n_j = 1$ for $1 \le j \le N$. Then pick up two patterns randomly, form a new (enlarged) pattern by aggregating events belonging to two patterns, and compute the error. If the error associated with the new pattern is smaller than the sum of errors from the previous two patterns, the new pattern is accepted. This procedure is repeated, until no two such patterns exist.

However, this simple classification scheme based on error minimization does not work well. The whole procedure depends critically on the choice of the error threshold ϵ and the error function itself. If ϵ were too small, most events remain isolated as its own pattern. On the other hand, if ϵ becomes too large, we frequently observe contradictory situation such as $Error(E_i, E_j) \le \epsilon$, $Error(E_j, E_k) \le \epsilon$ but $Error(E_i, E_k) > \epsilon$. It is also commonly observed that the classification results in a big cluster with few patterns.

Intrinsic noise in the observation data is an important factor to be considered in classification. If the noise level is very small, the classification based on the error function is reliable, provided the definition of the error function is sound and physically meaningful. On the contrary, if the data is severely contaminated by noise, the computed error is less reliable, and another way for compensating the noise would be required.

Another important point is the usefulness of the classification. If we set $\epsilon = 0$, then the classification is exact in the sense that the specification of the pattern class of an event is a complete description of the event. However, it yields too many categories. On the other extreme, if we use very large ϵ, most events belong to a single category. So the pattern seems to be an useful measure, as it can explain many events. However, the error is too large inside the category, rendering the classification useless. As a rough rule of thumb, we can say that $N = c \times \bar{N}_c$, where c is the number of categories, and $\bar{N}_c$ is the average number of the events belonging to the categories. There is a tradeoff between c and $\bar{N}_c$; if either of the two becomes too large, the classification becomes useless.

To cope with this problem, we adopt the formalism of thermal physics. Rather than minimizing the error, we introduce a free energy F, defined as

$$F = U - TS, \tag{5}$$

$$S = -\sum_{1 \le j \le c} p_c \log p_c, \quad p_c = n_c/N \tag{6}$$

where S is the entropy, a measure of the information associated with the classification. T is a control parameter representing a noise level in data, playing a role analogous to the temperature in thermodynamics. When T is small, U is dominant and error minimization becomes more important.

Table 1. Number of classified patterns and events classified at different noise parameter T.

T	Number of Patterns	Number of Events
0.25	11	29
0.26	16	41
0.27	19	55
0.28	24	79
0.29	27	111
0.30	34	143
0.31	40	194
0.32	43	224
0.33	45	253

When T is large, entropy term dominates, which prefers fewer number of patterns with larger number of events in patterns. Thus, the parameter T controls the amount of the distribution of the information between intra- and inter- pattern parts.

When $T = 0$, the results is the same for the error minimization scheme described previously. For nonzero values of T, similar procedure is performed. At first, every event is in its own pattern. Randomly pick two patterns, form a new pattern, and calculate the free energy difference. If it is negative, accept the new pattern. Repeat until no such patterns found.

Patterns at different T values may be computed independently, or may be continued from one T value to another. The continuation procedure works as follows. First, compute the pattern set at T_1. At a higher T_2, start the pattern classification based on the patterns computed at T_1, instead of starting from a pattern set made of individual events. With the continuation procedure, one can observe the evolution of the patterns with T. At lower T, patterns may have only a few number of events. With increasing T, more events are introduced to patterns, and patterns may be merged together. In our computation, we started at $T = 0.25$, and used continuation technique with increasing T.

Table 1 summarizes the results of the classification. At the lowest noise parameter $T = 0.25$, 11 nontrivial patterns with more than one event are found. Among the total of 807 events, 29 events are classified into one of the nontrivial patterns. Thus, on average, each pattern has three events. With increasing T, we can observe that more events are classified. The number of distinct patterns shows the increasing tendency, but not as evident as the case for the number of events. It is because different patterns at lower T may be merged into a single pattern at higher T. At $T = 0.33$, roughly 25% of the events are classified. With further increasing T, we do not observe a substantial improvement of the classification in terms of the number of events explained by patterns. We suspect that these 25% are the reasonable set of events that can be classified or identified as recurrent, whereas the rest 75% of the events are not classified. The coverage of the classification could be improved by a careful re-definition of an event, or the error function between events.

Figures 3 to 5 show examples of patterns at different noise parameter T. The average pattern is plotted as a line and the mean deviations as bars around the average. Note that at $T = 0.25$, we find a pattern possessing 8 events. With increasing T, more events are incorporated into the pattern and at $T = 0.30$, the pattern contains 13 events. Another pattern having 5 events at $T = 0.30$ is illustrated in Figure 5.

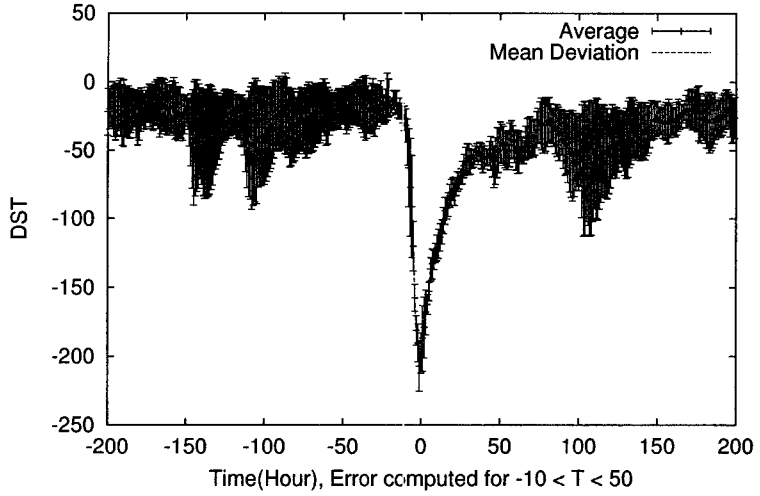

Figure 3. Pattern including Apr. 1, 1976 event at $T = 0.25$.

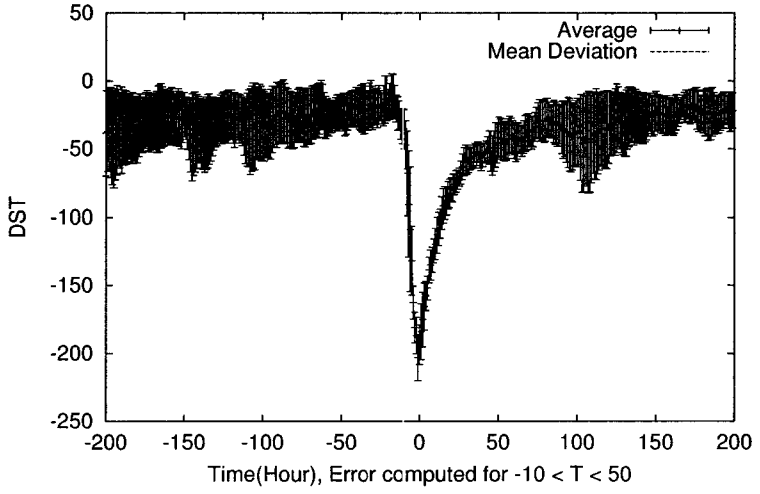

Figure 4. The same pattern as that in Figure 3 at an increased noise parameter $T = 0.30$

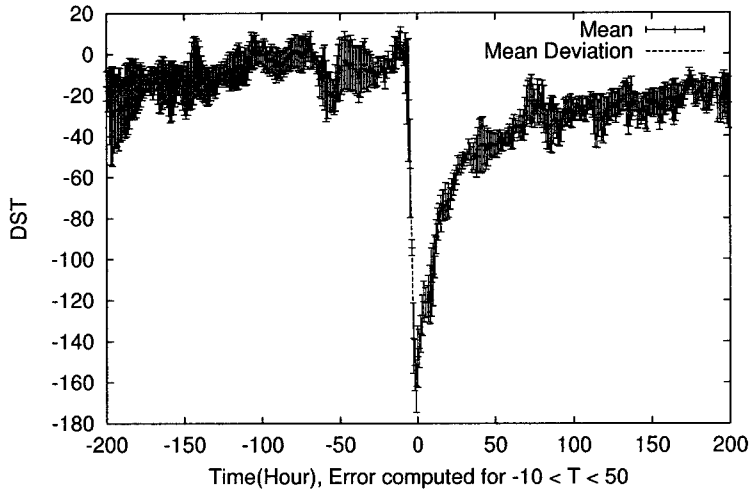

Figure 5. A different pattern containing Feb. 2, 1969 event at $T = 0.30$

3. DISCUSSION

We have demonstrated one approach for the classification of magnetic storms into a set of recurrent patterns. Interesting patterns have been discovered which a large number of the past storm events belong to.

In our approach, the classification is performed by seeking a categorization that minimizes the free energy. Thus, the classification results depend on two factors, the error and the entropy. By introducing free energy, we are essentially using "Maximum Entropy" or "Maximum Likelihood" estimation that is widely utilized in statistics. Roughly speaking, the principle tries to maximize (a priori) probability that a given sample is observed. Practically, it makes the categorization biased to assign a pattern into a class with larger number of population, thus maximizing a priori probability of observation. The noise parameter T is introduced to balance the Error and the Entropy. So, our classification results not only depend on the choice of the error function, but also on the estimation of the degree of noise present in the data. In order to find a physically relevant classification result, one should change the parameter T as well as the error function and the characteristic function. The procedure can be performed as trial-and-error basis, or using a systematic search over a set of candidate functions and a range of T values.

Although the error in this work was defined as a squared sum of the value differences for simplicity, there can be many other choices of the error function. For example, some authors noticed that storm patterns look quite alike after normalization where all the values are divided by the absolute value of the Dst minimum during the storm. In that case, re-scaled error function, in which the error is computed after re-scaling the events, would be more appropriate as one can focus on the overall shape of the events.

The number of data points (hours) used for the Error computation can also affect the results. Typical development of a storm is usually divided into pre-storm, main phase, and recovery phase. One can choose the data points so that only the main phase differences contribute to the error. Or

one can also sum up all the differences for the whole storm phase. One may further insist upon some kind of preprocessing to reduce noise, such as smoothing.

In the classification procedure, we randomly picked two events and computed free energy difference. Therefore, the classification can be affected by the sequence of random numbers generated. To amend this problem, a more sophisticated statistical approach is required. At fixed T, the classification is repeated using different random sequence, and the probability $p(i, j)$ that event i and event j belong to the same pattern is computed. If it were larger than a certain threshold, they are accepted to belong to the same category. We used binary comparison, i.e., comparing two patterns at a time. In some cases, a merge of three patterns may reduce the free energy, whereas the merges between any pair of them increase the free energy. Therefore it is also necessary to consider the possibility of tertiary and higher order test in classification.

A choice of an error function cannot be fully justified without a proof of physical relevance. We implicitly assume that the storm pattern is a deterministic output of initial settings, including solar wind and geomagnetic environment before the onset of the storm. A suitable error function should report small errors for storms developed from similar initial conditions, and large errors otherwise. As the geomagnetic activity is controlled by solar wind conditions (Nishida 1983), one may reasonably focus on the input solar wind conditions as a good proxy of the full initial conditions before a magnetic storm. If we find the input patterns in the solar wind conditions, which are closely related to the output patterns in Dst index, the relation between the input solar wind patterns and output Dst index may be used for space weather forecasting. It should be an important next step of this study.

ACKNOWLEDGEMENTS: This work was supported by Korea Research Foundation Grant (KRF-2002-037-C00014).

REFERENCES

Burton, R. K., McPherron, R. L., & Russel, C. T. 1975, JGR, 80, 4204
Nishida, A. 1983, Space Sci. Rev., 34, 185
Su, S. Y., & Konradi, A. 1975, JGR, 80, 195