THE ASTROPHYSICAL JOURNAL, 566:202–209, 2002 February 10

CLASSIFYING GAMMA-RAY BURSTS USING SELF-ORGANIZING MAPS

H. J. RAJANIEMI[1,2] AND P. MÄHÖNEN[1]

Received 2001 June 21; accepted 2001 October 9

ABSTRACT

Using the self-organizing map (SOM) algorithm developed by Kohonen in 1982, we investigate the three gamma-ray burst classes discovered by Mukherjee et al. in 1998 and classify the bursts in the current BATSE catalog. The validity of the classification and the isotropy of the burst classes is examined using well-known statistical methods. We are able to show by this independent unsupervised classifier that two previously observed burst classes are indeed present in BATSE data. We also support the 2000 conclusion of Hakkila et al. that there is no strong statistical signal for the existence of the third class, which might actually be due to an instrumental bias. This result is obtained through our independent analysis method, which is not related to the 2000 analysis of Hakkila et al.

Subject headings: gamma rays: bursts — methods: analytical — methods: data analysis — methods: statistical

1. INTRODUCTION

Although 3 decades have passed since their discovery in 1969 by the US *Vela* satellite (Strong, Klebesadel, & Olson 1974), gamma-ray bursts (GRBs) remain a puzzling astronomical mystery. Very little is known about the fundamental physical mechanisms behind these phenomena, although a bewildering variety of models exist in the literature. GRBs exhibit great morphological diversity and complex temporal behavior: attempts to divide them into classes according to their light curves have been unsuccessful. Studies based on GRB bulk properties (such as burst duration and spectral hardness) have proved to be more fruitful, and the discovery of distinct classes in the GRB population might lead to some new astrophysical insight.

Evidence for duration bimodality among GRBs was found as early as 1974 (Strong et al. 1974), and the division of bursts into long-soft and short-hard classes remains the most widely accepted taxonomy. Kouveliotou et al. (1993) identified two types of bursts in the first BATSE[3] catalog using the burst duration parameter T_{90} (the time it takes for 90% of the total burst flux to arrive), noting that there is a dividing line around $T_{90} \sim 2$ s: bursts longer than this are significantly softer and brighter than the rest. This is at odds with the predictions of cosmological models, which indicate that long faint bursts at high redshift should also be softer than bright ones at low redshift (Hakkila et al. 2000). Otherwise, the long-soft (henceforth class 1) and the short-hard (class 2) do not differ markedly, although it has been pointed out that the groups may have different Galactic latitude distributions (Belli 1997) and fluence distributions (Katz & Canel 1996).

It has been argued that there might exist a third class of GRBs, an "intermediate" class 3 consisting of soft bursts with durations and intensities falling between those of class 1 and class 2 bursts. Horváth (1998) finds that a three-Gaussian fit of log T_{90} is significantly more probable than a two-Gaussian fit. In addition, the detailed statistical clustering analysis of Mukherjee et al. (1998) seems to favor three classes, with support from Bayesian likelihood estimates.

Many other possible GRB subclasses have been suggested, based on properties such as differing short-timescale variability (Lamb, Graziani, & Smith 1993) and different ratios of total fluence to greater than 300 keV fluence (Pendleton et al. 1997). Recently, Cline et al. (1999) have proposed that very short (shorter than 100 ms) bursts form a separate subclass. Nevertheless, the long-soft/short-hard/intermediate classification scheme remains the main focus of attention.

Although data from the current BATSE catalog yields itself to statistical methods such as those applied by Mukherjee et al. (1998), one might ask if there exist alternative techniques for the complex task of GRB classification. Artificial intelligence–based (AI-based) pattern recognition algorithms are one possible candidate: automated linear classification of vector data into a given number (or an arbitrary number) of classes is a well-established problem in the field of machine learning. Several varieties of AI-based classifiers exist. Usually, one makes the distinction between supervised and unsupervised classifiers: the former are trained with data for which the classification is known and then used to classify raw data, while the latter attempt to find the best-fitting class structure in the input data by using some measure of merit (Vesanto & Alhoniemi 2000; Hertz, Krogh, & Palmer 1991).

Baumgart (1994) generated a taxonomy of GRB data collected by the *Pioneer Venus Orbiter* (*PVO*) spacecraft using an unsupervised neural network. There were a total of 26 variables for each of the 99 bursts in the data set, including both bulk and temporal properties and some more exotic ones, such as the fractal dimension of the signal. Baumgart's network found two or three distinct classes in the data set, but given the small size of the *PVO* catalog and the large number of parameters used, it is difficult to compare his work with later results.

A more extensive study of gamma-ray burst classes using AI techniques was conducted recently by Hakkila et al. (2000): the supervised decision tree classifier C4.5 was trained with BATSE 3B data classified by Mukherjee et al. (1998), and the program generated a set of rules delineating

[1] University of Oulu, Department of Physical Sciences, Linnanmaa, FI-90570 Oulu, Finland.
[2] University of Edinburgh, Department of Mathematics, Edinburgh, EH9 3JZ, UK.
[3] The Burst and Transient Source Experiment in NASA's *Compton Gamma Ray Observatory* (*CGRO*).

the class properties. Hakkila et al. also identified a new bias in the BATSE instrument, suggesting that the third class may be an instrumental effect and not a true separate source population, and demonstrated that classes 1 and 2 can be distinguished using only the Band spectral parameters α, β, and $E_{\rm peak}$. This paper is largely inspired by the work of Hakkila et al. (2000). Instead of using a decision tree classifier, we employ the self-organizing map (SOM) algorithm developed by Kohonen (1982). We demonstrate that the SOM is a viable tool in GRB classification and use it to explore the properties of data in the current BATSE catalog, examining the validity of the classification generated by SOM using statistical techniques. Finally, the isotropy of the GRB classes is investigated.

2. DESCRIPTION OF THE SOM ALGORITHM

Since their introduction in 1984, Kohonen's SOMs have become a very popular tool for a wide range of applications, in particular the clustering and visualization of high-dimensional data sets. The literature surrounding the subject is vast, and a recent survey (Kohonen 1997) lists more than 3000 references. In spite of its popularity and flexibility, the SOM has not yet been widely used in astronomical applications. The suitable and more general treatment on neural networks, including an SOM algorithm for astronomers, is written by Haykin (1999).

Many variants of SOM exist, but the basic philosophy behind them all is the same: unlike many other types of neural networks, the SOM tries to preserve the *topology* of the input data. It defines a mapping from the input data space ($\boldsymbol{R}^n$ for convenience) onto a regular, usually two-dimensional lattice of map units (also called "neurons"), which are connected to adjacent ones with some neighborhood relation. With each unit i, one associates a prototype vector m_i. These prototype vectors are also referred to as "weights" (w_i) in the neural networks literature, since they are used by the updating algorithm of the neural network. The dimensionality of the SOM is determined by the number of neurons M in the map (in the case of a two-dimensional map, this is usually written as $M = N \times N$) and by the number of input parameters (also often referred to as input neurons). The prototype vectors can be initialized randomly or linearly along a two-dimensional subspace spanned by the two principal eigenvectors of the input data. One can get a geometrical idea about the SOM by looking at Figure 1, which represents a 9×9 network. The prototype vectors m_i are associated with each neuron in the SOM. The update algorithm maps input vectors (in our case, three-dimensional vectors) into this topological map. In an extreme case, one could consider 81 separate classes, each associated with a different neuron. As is demonstrated below, in practice the true number of clusters is lower, determined by sufficient dissimilarity between groups of neurons.

During training, each input data instance x (which is a vector $x \in \boldsymbol{R}^n$: its dimension defines the number of input variables) is mapped onto the best-matching unit. This is defined to be the neuron b (from the set of $N \times N$ neurons) with prototype vector m_b closest to x. The closest match is found simply by computing the following statistic for each neuron in the map:

$$D_i = \min_i \left(\| x - m_i \| \right) , \tag{1}$$

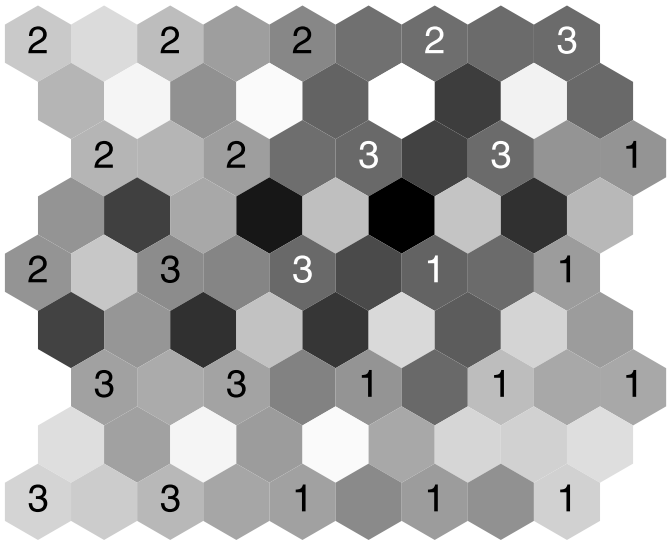

FIG. 1.—U-matrix of an SOM trained with 150 random bursts from the BATSE 3B catalog. Shades of gray indicate distances between map units: darker shades represent large distances.

where i runs over all neurons in the two-dimensional SOM network. This means that we simply calculate the difference between the input vector and each prototype vector m_i that is associated with the neuron i. The closest match, or "winner," is the one with the smallest difference. The index of the winning neuron is often denoted by b, i.e., $\| x - m_b \|$ is the smallest difference during each training step.

Instead of the Euclidean distance, other well-defined metrics could be used. An often-used alternative is a dot product $x \cdot m_i$, in which case the winning neuron is not the neuron with the minimal value of the dot product, but instead the neuron with the maximal value. We stress that in each instance (or training step), the winning neuron's prototype vector (weight) is the one most closely resembling the given input vector x. Next, the best-matching unit and the units in its topological neighborhood are shifted closer to the input vector. The update rule is

$$m_i(t+1) = m_i(t) + \alpha(t) h_{bi}(t) [x - m_i(t)] , \tag{2}$$

where t is a discrete time coordinate, h_{bi} is the so-called neighborhood kernel, a function describing the shape of the unit neighborhoods, and $\alpha(t)$ is a scalar-valued "learning rate" that is an adaptation coefficient that represents the learning rate of the map and decreases monotonically with time. The above steps are repeated iteratively until one obtains a sufficiently accurate map. What is meant by "sufficient accuracy" is naturally different depending on what qualities of the input data one wishes to examine. As a rule of thumb, the "average quantization error," i.e., the mean of $\| x - m_b \|$, is a useful measure of performance.

Intuitively, one can think of the SOM as an elastic surface that is fitted over the data points: the neighborhood kernel then represents the stiffness of the surface. A widely applied kernel function is the Gaussian kernel:

$$h_{bi}(t) = \exp \left[- \frac{\| r_b - r_i \|^2}{2\sigma^2(t)} \right] , \tag{3}$$

where the parameter $\sigma(t)$ describes the width of the kernel, and, like α, is chosen to decrease monotonically with time. The variables r_b and r_i are the radius vectors of neurons b and i, respectively, and t is again the discrete time coordi-

nate. Typically, parameters $\alpha(t_0)$ and $\sigma(t_0)$ are chosen based on the size of the SOM network and by experimenting with data to find a reasonable learning rate.

The network is trained by selecting a representative set of input vectors. The SOM is initialized, and after that, training vectors are fed into the SOM by the above algorithm (see more details in Kohonen 1997; Mähönen & Hakala 1995). Once the SOM has been successfully trained, the prototype vectors (or weights) contain the cluster means for the classification. The subsequent test set (or new data) can be simply classified by giving an input vector to the SOM and finding the winning neuron that represents the cluster to which the input vector belongs.

The SOM is closely related to the classical K-means clustering; the two algorithms are, in fact, identical if one sets the SOM neighborhood function to zero. By using a number of map units roughly equal to the expected number of clusters, the SOM can be used to classify a given data set into distinct groups.

However, the SOM is not merely a clustering technique: it can also be used to visualize the input data set, since it attempts to represent the topological ordering of the input data in the ordering of the map unit reference vectors. In this case, it is common practice to use an SOM with a relatively large number of map units, up to a multiple of the number of input data vectors used in training. One can then examine the structure of the data set and identify potential clusters visually. A widely used technique is the so-called U-matrix: distances of each map unit to its immediate neighbors are computed and used as a measure of dissimilarity of the map units. The distances can be displayed as heights on the map landscape, for instance, or as varying shades of gray. Clusters on the map then show up as "valleys" separated by "hills" or "walls" (Ultsch & Siemon 1990).

Both of these approaches are used in the following section. We test the clustering properties of the SOM using the BATSE 3B data classified by Mukherjee et al. (1998), classify the bursts in the current BATSE catalog, and visualize the GRB data using the U-matrix.

We stress that for our work, it is important that the SOM is an *unsupervised* clustering algorithm. Specifically, this means that there is no need to classify (cluster) training data by hand or other methods. This is an important difference when compared to some other neural networks, such as back-propagation networks, that require classified training data to teach them. However, it is very important that the training data is large enough and a representative set of instances of the whole phenomena under study. Therefore, "unsupervised learning" refers in this paper to the neural networking and clustering algorithm property that classification is done automatically, without any human observer preclassification. However, naturally, observers must provide a good and "representative" data set for this method that ultimately requires some judgement from an observational point of view.

All SOM simulations were conducted using the SOM_PAK software package.[4]

3. CLASSIFYING GRBs IN THE CURRENT BATSE CATALOG

Mukherjee et al. (1998) used a variety of hierarchical clustering techniques to study the bursts in the BATSE 3B catalog. After eliminating noise-dominated burst attributes, they were left with a six-dimensional parameter space (logarithmically normalized total fluence $S_{\rm tot}$, peak flux P_{256}, durations T_{90} and T_{50}, and hardness ratios HR_{321} and HR_{32}). A multivariate analysis of the data was then conducted using standard nonparametric hierarchical clustering analysis and a parametric maximum-likelihood model-based clustering procedure. All methods employed indicated that three subclasses were present, and Bayesian analysis supported this result (giving three as the most probable number of classes), based on the assumption that the burst attributes were drawn from a mixture of multivariate Gaussian distributions. However, the third subclass appeared only weakly in some cases, and exact class assignments depended on the clustering algorithms used. Nevertheless, we use the classification of Mukherjee et al. (1998) as a starting point.

3.1. *Two or Three GRB Classes?*

Hakkila et al. (2000) identify a fluence duration bias affecting BATSE measurements and show that it can cause burst durations to be underestimated. They argue that this bias is responsible for some class 1 bursts taking on class 3 characteristics (i.e., some class 1 bursts appear shorter and fainter than expected). Thus, the third subclass might not represent a separate source population, but simply distant class 1 bursts or bursts belonging to the short, soft end of the class 1 distribution. The fact that Hakkila et al. (2000) were unable to distinguish class 3 bursts from class 1 bursts based on the Band spectral parameters α, β, and $E_{\rm peak}$ alone, even though classes 1 and 2 are spectrally distinct, provides further evidence for this hypothesis.

We decided to test the clustering properties of the SOM against the classification of Mukherjee et al. (1998), and then examine the validity of the aforementioned hypothesis by reassigning class 3 bursts to class 1. For convenience, we refer to the two classes obtained by assuming that the bias proposed by Hakkila et al. (2000) is correct as class 1b and class 2b.

3.2. *Choice of Parameters*

One important question to be addressed before training the SOM was the choice of parameters used. The principal component and factor analysis conducted by Bagoly et al. (1998) on data from the BATSE 3B catalog indicates that only three variables are needed to characterize the relationships between the bursts in the database, namely the weighted fluence, burst duration, and flux in the highest energy bin. Following Hakkila et al. (2000), we chose $\log T_{90}$, $\log \mathrm{HR}_{321}$, and $\log S$, where S is the sum of channel 2 and 3 fluxes. Experimentation with a larger number of parameters showed no improvement in classification accuracy.

3.3. *Classification Methodology and Results*

We used an SOM with 5×5 neurons, randomly initialized weight vectors, and hexagonal topology. One hundred training sets were generated by randomly picking 50 bursts from each of the classes. Each of the resulting sets was then used to train 100 randomly initialized maps, and the map with the smallest average quantization error was chosen for further examination. The map was then labeled using the training data set for map calibration. Each map unit was assigned to one of the three classes by simple majority

[4] Available on the Web at http://www.cis.hut.fi/research/som_pak.

TABLE 1

BURST CLASS PROPERTIES: MEANS AND STANDARD DEVIATIONS

Variable	Class 1	Class 2	Class 3	Class 1b	Class 2b
log T_{90}	1.61 ± 0.38	-0.36 ± 0.46	0.75 ± 0.29	1.47 ± 0.48	-0.32 ± 0.49
log S.....	-5.37 ± 1.09	-6.56 ± 1.55	-6.27 ± 1.11	-5.52 ± 1.15	-6.57 ± 1.50
log HR_{321}	0.16 ± 0.25	0.51 ± 0.29	0.07 ± 0.41	0.14 ± 0.29	0.51 ± 0.29
Number of bursts N_B.....	1236	507	294	1499	538

voting; i.e., each unit was given the same class label as the majority of bursts assigned to that unit during training.

Next, the BATSE 3B bursts were reclassified using the map simply by assigning each burst to the same class as its best-matching unit. The best map correctly (i.e., in agreement with Mukherjee et al. 1998) classified 93% of the bursts (96.9% agreement for class 1, 94.5% for class 2, and 72.6% for class 3). The U-matrix of the best map is shown in Figure 1: classes 1 and 2 are visually distinct, even in a map this small, but class 3 is not. Notably, the map misclassifies class 3 bursts far more often than class 1 or class 2 bursts. Examination shows that the map has a tendency to assign class 3 bursts to class 1 (86.7% of the time).

Testing the bias hypothesis of Hakkila et al. (2000), we then assigned the class 3 bursts in the catalog to class 1 and repeated the above procedure. This time, the resulting classification accuracy was 96.9% (96.6% agreement with class 1b and 97.5% for class 2b). The two classes are very clearly distinct on the resulting map, as Figure 2 shows.

The existence of class 3 as a separate subclass can be further examined by visualizing the entire current BATSE catalog using an SOM. In this case, a map with 10×10 was used, with the current catalog as the training set. The 3B bursts classified by Mukherjee et al. (1998) were used for map calibration. Figures 3 and 4 show the U-matrix of the resulting map, in gray scale and as a three-dimensional landscape plot, respectively. Again, classes 1 and 2 form distinct clusters separated by a "mountain range," indicating that the two clusters are separate in the parameter space. Map units belonging to class 3 are mostly placed on the border between classes 1 and 2.

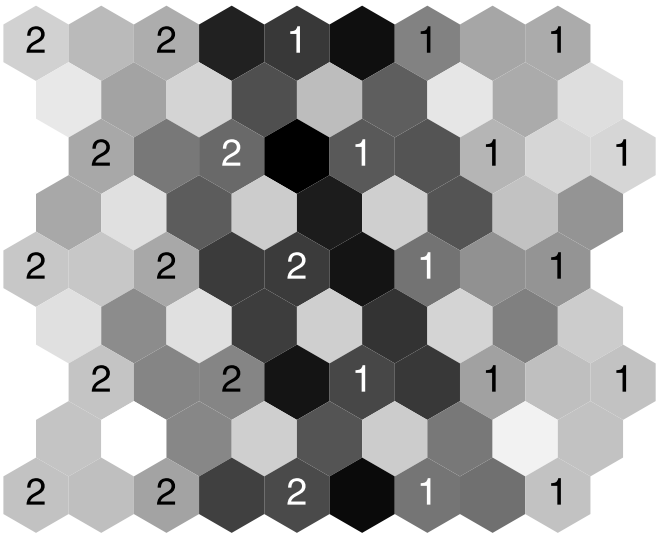

FIG. 2.—U-matrix of an SOM trained with 100 random GRBs from classes 1b and 2b.

One should, of course, exercise caution when interpreting these results, but the fact that the SOM is unable to delineate the third class clearly does raise some qualitative doubt as to its existence as a separate source population.

As a final step in the analysis, we trained an SOM using the entire BATSE 3B data set (797 bursts) as training and calibration data and classified the 2041 bursts in the current BATSE catalog: again, first assuming that there are three separate classes, and then assuming that the bias hypothesis holds and assigning class 3 bursts to class 1. Table 1 summarizes some of the properties of the classes generated by the SOM, and Figures 5 and 6 demonstrate the evident bimodality of the T_{90} parameter.

Figures 7 and 8 show the sky distributions of the burst classes. There is no spatial correlation related to the classification.

3.4. *Testing the Validity of the Classification*

Assuming that the burst properties are drawn from a mixture of multivariate Gaussian distributions, it is possible to evaluate the statistical significance of our classification using multivariate analysis of variance (MANOVA) tests. Mukherjee et al. (1998) outline some of these methods and apply them to their own classification scheme: further details can be found in Anderson (1984). Table 2 gives three statistics testing the null hypothesis that classes 1, 2, and 3 (and likewise, classes 1b and 2b) are drawn from the same distribution and have the same mean: Wilks's lambda Λ, Pillai's trace V, and Hotelling-Lawley's trace U. Roughly speaking, these statistics measure the offset of cluster means from each other. They can be computed by defining two matrices of sums of squares and cross products:

$$\boldsymbol{B} = \sum_{l=1}^{g} n_l (\overline{x}_l - \overline{x})(\overline{x}_l - \overline{x})^T , \tag{4}$$

$$\boldsymbol{W} = \sum_{l=1}^{g} \sum_{j=1}^{nl} (x_{lj} - \overline{x}_l)(x_{lj} - \overline{x}_l)^T . \tag{5}$$

Here g is the number of clusters, n_l is the number of vectors in the lth cluster, $\overline{x}_l$ is the centroid of the lth cluster, and x_{jl} denotes the jth data vector in the lth cluster.

The terms Λ, V, and U are now calculated from the matrices $\boldsymbol{B}$ and $\boldsymbol{W}$ as

$$\Lambda = \det(\boldsymbol{W})/\det(\boldsymbol{B} + \boldsymbol{W}) , \tag{6}$$

$$V = \mathrm{trace}[\boldsymbol{B}(\boldsymbol{B} + \boldsymbol{W})^{-1}] , \tag{7}$$

$$U = \mathrm{trace}(\boldsymbol{W}^{-1}\boldsymbol{B}) . \tag{8}$$

The asymptotic distributions of the three statistics are known. In particular, the so-called F approximation derived from Wilks's Λ follows the F distribution, allowing us to compute the statistical significance of the obtained values. In all cases, the value of F is very high, indicating

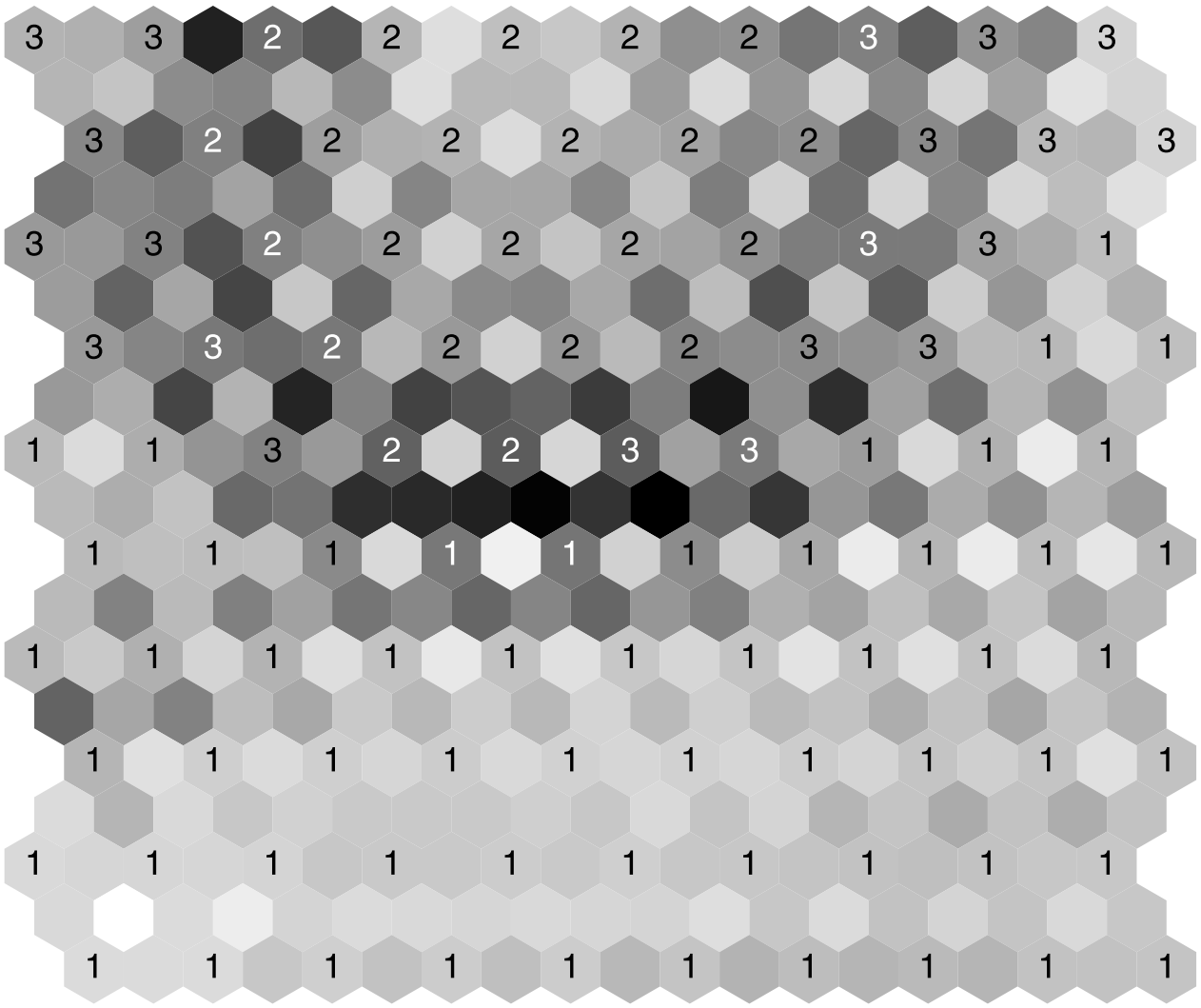

FIG. 3.—U-matrix of an SOM trained with the bursts in the current BATSE catalog

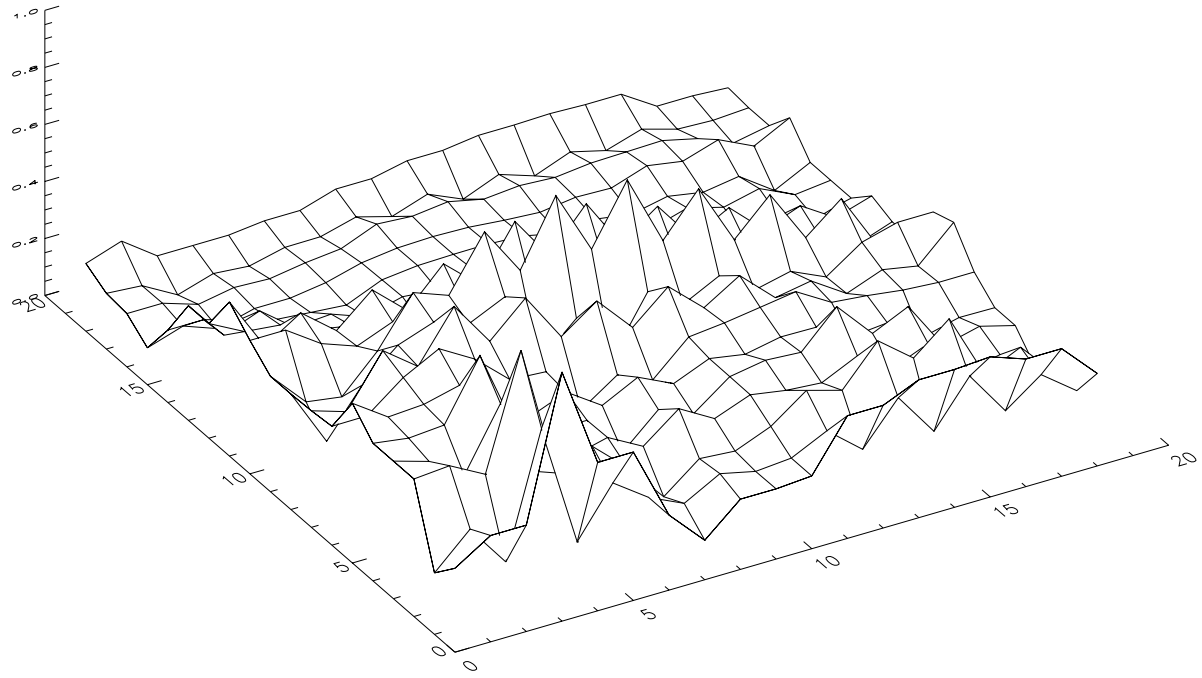

FIG. 4.—Landscape plot of the U-matrix of the map trained with the current BATSE catalog bursts. Classes 1 and 2 are separated by a "mountain range."

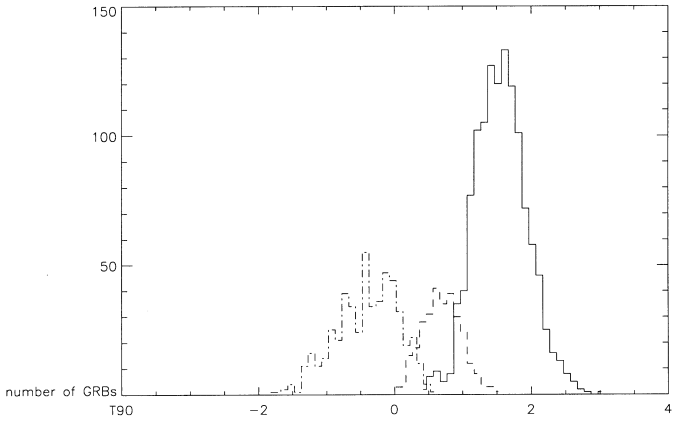

FIG. 5.—Plot of log T_{90} histograms for classes 1, 2, and 3

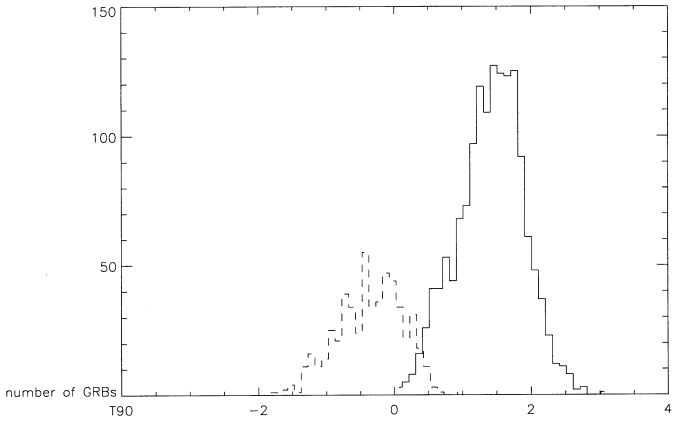

FIG. 6.—Plot of log T_{90} histograms for classes 1b and 2b

that it is very unlikely that the GRB subgroups are drawn from distributions that have the same mean.

It should be noted that it is dangerous to draw too far reaching conclusions from the results of these tests, since the computation of the above statistics is based on the assumption that the data is drawn from a mixture of multivariate Gaussian distributions. Nevertheless, the tests serve as a quantitative demonstration that different GRB classes do exist in the current BATSE catalog, at even higher levels of statistical significance than those reported by Mukherjee et al. (1998) for the BATSE 3B data.

3.5. *Isotropy of the GRB Classes*

The remarkable isotropy of the GRB distribution in the sky has been an important constraint on theories about GRB origins, indicating that GRBs must lie at cosmological distances. Even so, it is possible that although the bulk of GRBs are isotropic, some subclass might display aniso-

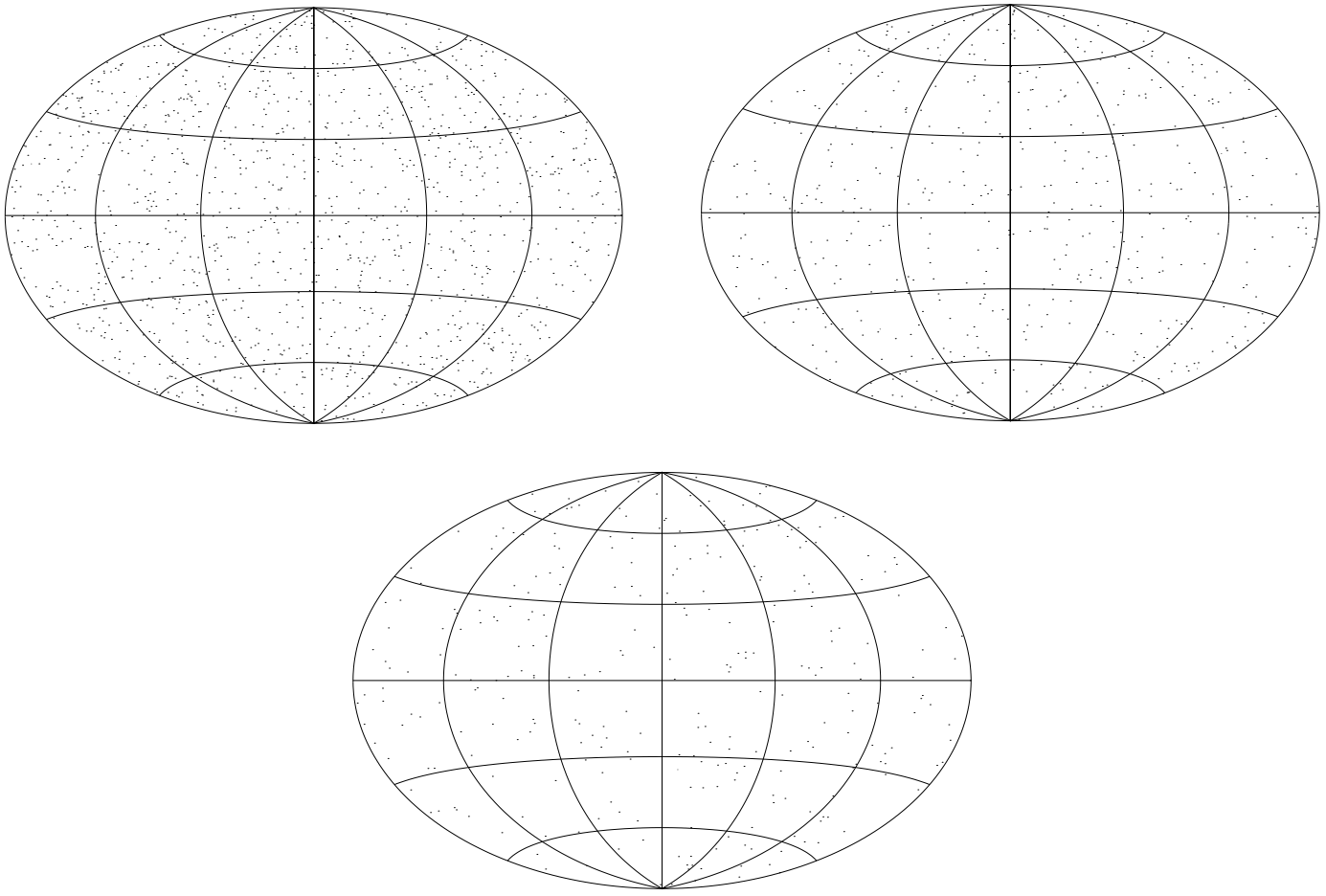

FIG. 7.—Sky distributions of GRB classes 1, 2, and 3 using the Hammer-Aitoff projection

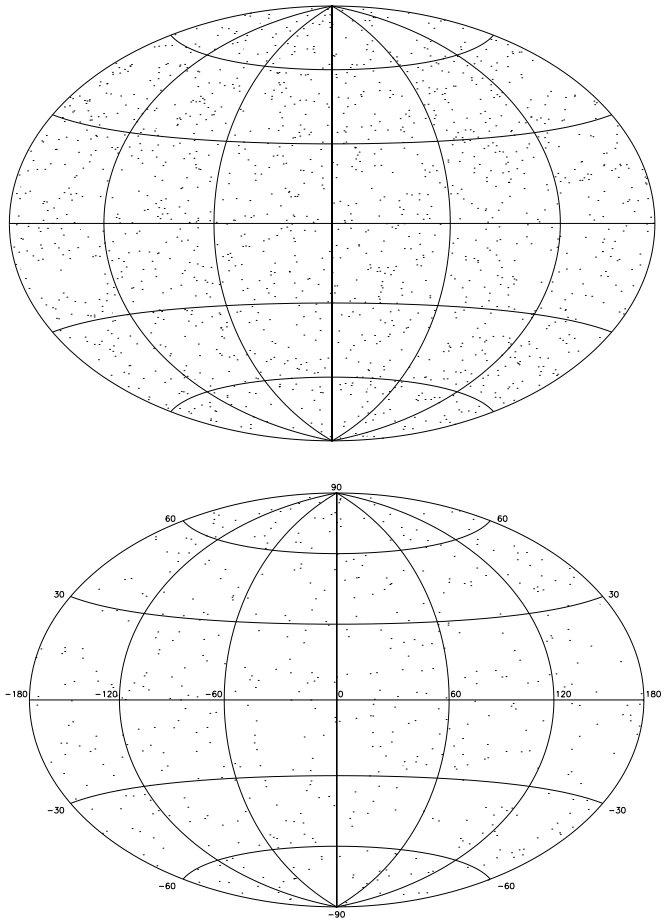

FIG. 8.—Sky distributions of classes 1b and 2b

tropy, indicating galactic origin. For instance, Cline et al. (1999) report a nonisotropic angular distribution for very short GRBs. In addition, Balázs, Mészáros, & Horváth (1998), Mészáros et al. (2000a), and Mészáros, Bagoly, & Vavrek (2000b) report a significant anisotropy in the sky distribution of the intermediate subclass 3.

We applied four standard statistical isotropy tests to the GRB subsets generated by the SOM. A review of standard isotropy tests can be found in Briggs (1993); these tests are applied to various subsets of the BATSE 3B catalog in Briggs et al. (1996). Results are summarized in Tables 3 and 4. The expected values for the statistics corrected with the BATSE sky exposure function were calculated by Briggs et al. (1996).

We find that our results are consistent with isotropy; all values lie within one standard deviation of the expected values. The only exceptions, $\mathscr{B} = 10.616$ for class 2 and $\mathscr{B} = 11.860$ for class 2b, have very low statistical significance. The dipole and quadrupole moments are very close to the expected values in all cases, indicating a high level of isotropy in all subclasses.

4. CONCLUSIONS

We have demonstrated that it is possible to use Kohonen's SOM algorithm as a tool for GRB classification and data mining. Statistical techniques have been applied to examine the validity of the GRB taxonomy generated by the SOM, and the isotropy of the burst classes has been examined. SOM visualization of the current BATSE catalog data gives some qualitative and quantitative support to the hypothesis that the existence of the intermediate GRB class is due to an instrumental bias, but the evidence remains inconclusive. The SOM proves to be a fast

TABLE 2

MANOVA TEST RESULTS

Classes	Wilks's Λ	Pillai's Trace	Hotelling-Lawley	F Approximation	Probability of Equal Means
1, 2, 3.....	0.542	0.499	0.771	243.274	$<10^{-6}$
1, 2.....	0.570	0.430	0.753	437.560	$<10^{-6}$
1, 3.....	0.584	0.416	0.712	362.285	$<10^{-6}$
2, 3.....	0.657	0.342	0.521	138.978	$<10^{-6}$
1b, 2b.....	0.568	0.433	0.762	517.489	$<10^{-6}$

TABLE 3

ISOTROPY PROPERTIES OF GRB CLASSES

Statistic	Class 1	Class 2	Class 3	Expected Value for Isotropy
Dipole moment $\langle\cos\theta\rangle$.....	−0.0201	−0.0046	0.0086	−0.013
Quadrupole moment $\langle\sin^2 b - 1/3\rangle$.....	−0.00385	0.00713	−0.0821	−0.005
Rayleigh-Watson $\mathscr{W}$	1.324	0.7442	1.611	$3 + 0.002N_B$
Binghham $\mathscr{B}$	5.2626	10.6157	8.28525	$5 + 0.0077N_B$

TABLE 4

ISOTROPY PROPERTIES FOR CLASSES 1B AND 2B

Statistic	Class 1b	Class 2b	Expected Value for Isotropy
Dipole moment $\langle\cos\theta\rangle$.....	−0.01499	−0.00418	−0.013
Quadrupole moment $\langle\sin^2 b - 1/3\rangle$.....	−0.00368	0.003624	−0.005
Rayleigh-Watson $\mathscr{W}$	1.9069	0.6514	$3 + 0.002N_B$
Binghham $\mathscr{B}$	8.285	11.860	$5 + 0.0077N_B$

and reliable GRB classifier, and its flexible nature makes it easy to extend the methods outlined here to future GRB catalogs. Because an SOM is an unsupervised and topology-preserving classification algorithm, it is a markedly different and independent method compared to previous analysis tools used for the classification of GRBs. Hence, the results obtained through SOM analysis are giving important and independent corroboration to previous results. SOM classification of GRBs based on GRB light curves or Band spectral parameters are two possible future directions for our research.

This work has been supported in part by the University of Oulu through Science Council Funding and the Academy of Finland. H. J. R. acknowledges research student funding through the cosmology group and P. M. acknowledges the faculty award from the University of Oulu.

REFERENCES

Anderson, T. W. 1984, An Introduction to Multivariate Statistical Analysis (New York: Wiley)
Bagoly, Z., et al. 1998, ApJ, 498, 342
Balázs, L. G., Mészáros, A., & Horváth, I. 1998, A&A, 339, 1
Baumgart, C. W. 1994, Proc. SPIE, 2243, 553
Belli, B. M. 1997, ApJ, 479, L31
Briggs, M. S. 1993, ApJ, 407, 126
Briggs, M. S., et al. 1996, ApJ, 459, 40
Cline, D. B., et al. 1999, ApJ, 527, 827
Hakkila, J., et al. 2000, ApJ, 538, 165
Haykin, S. 1999, Neural Networks: A Comprehensive Foundation (2d ed.; Englewood Cliffs: Prentice-Hall)
Hertz, J., Krogh, A., & Palmer, R. 1991, Introduction to the Theory of Neural Computation (Redwood City: Addison-Wesley)
Horváth, I. 1998, ApJ, 508, 757
Katz, J. I., & Canel, L. M. 1996, ApJ, 471, 915
Kohonen, T. 1982, Biological Cybernetics, 43, 59
———. 1997, Self-Organizing Maps (New York: Springer)
Kouveliotou, C., et al. 1993, ApJ, 413, L101
Lamb, D. Q., Graziani, C., & Smith, I. A. 1993, ApJ, 413, L11
Mähönen, P., & Hakala, P. 1995, ApJ, 452, L77
Mészáros, A., Bagoly, Z., Horváth, I., Balázs, L. G., & Vavrek, R. 2000a, ApJ, 539, 98
Mészáros, A., Bagoly, Z., & Vavrek, R. 2000b, A&A, 354, 1
Mukherjee, S., et al. 1998, ApJ, 508, 314
Pendleton, G. N., et al. 1997, ApJ, 489, 175
Strong, I. B., Klebesadel, R. W., & Olson, R. A. 1974, ApJ, 188, L1
Ultsch, A., & Siemon, H. P. 1990, in Proc. Int. Neural Networks Conf. (Dordrecht: Kluwer), 305
Vesanto, J., & Alhoniemi, E. 2000, IEEE Transactions on Neural Networks, 11, 586