A TECHNIQUE FOR AUTOMATIC CLASSIFICATION OF METEOR TRAILS AND OTHER PROPA. (U) SIGNATRON INC LEXINGTON MA
J A WEITZEN ET AL. SEP 86 A447-13 RADC-TR-86-117

UNCLASSIFIED F19628-84-C-0117
A TECHNIQUE FOR AUTOMATIC CLASSIFICATION OF METEOR TRAILS AND OTHER PROPAGATION MECHANISMS FOR THE AIR FORCE HIGH LATITUDE METEOR BURST TEST BED

Signatron, Inc.

Jay A. Weitzen and Sylvia Tolman

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700
**Title:** A Technique for Automatic Classification of Meteor Trails and Other Propagation Mechanisms for the Air Force High Latitude Meteor Burst Test Bed

**Personal Authors:** Jay A. Weltzen, Sylvia Tolman

**Abstract:** This report describes a rule-driven artificial intelligence technique for automating the identification of propagation mechanisms and classification of meteor trail types on the USAF high latitude meteor burst test bed. The procedure reduces the time required to process data from one month of test bed operation from 2 staff months to about 12 hours. It is designed to emulate the procedures used by a human operator with the additional advantage of being able to use previous and future data records in the decision process. The performance of the automatic classification procedure is comparable in terms of correct decisions to human classifiers.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>SECTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 AN AUTOMATIC CLASSIFICATION PROCEDURE</td>
<td>3</td>
</tr>
<tr>
<td>2.1 PROPAGATION MODEL DEFINITION</td>
<td>3</td>
</tr>
<tr>
<td>3 THE AUTOMATIC CLASSIFICATION HEURISTICS</td>
<td>7</td>
</tr>
<tr>
<td>3.1 IDENTIFYING THE DOMINANT PROPAGATION MECHANISM</td>
<td>7</td>
</tr>
<tr>
<td>3.2 METEOR TRAIL CLASSIFICATION</td>
<td>10</td>
</tr>
<tr>
<td>3.3 EVALUATING THE ACCURACY OF THE CLASSIFIER</td>
<td>12</td>
</tr>
<tr>
<td>4 CONCLUSIONS</td>
<td>15</td>
</tr>
<tr>
<td>5 ACKNOWLEDGMENT</td>
<td>16</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>17</td>
</tr>
<tr>
<td>FIGURE</td>
<td>PAGE</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
</tr>
<tr>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>12</td>
<td>29</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>14</td>
<td>31</td>
</tr>
<tr>
<td>15</td>
<td>32</td>
</tr>
<tr>
<td>16</td>
<td>33</td>
</tr>
<tr>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>18</td>
<td>35</td>
</tr>
<tr>
<td>19</td>
<td>36</td>
</tr>
</tbody>
</table>
SECTION 1
INTRODUCTION

In order to evaluate the effects of high latitude on meteor burst communication, the U.S. Air Force has established a High Latitude Meteor Burst Test Bed between Sondrestrom AB, Greenland and Thule AB, Greenland. The 1260 km link with transmitter at Sondrestrom AB and receiver at Thule AB operates continuously cycling between four operating frequencies (45, 65, 104 and 147 MHz) every thirty minutes. Four-second data records of the envelope of the received CW signal are collected whenever the received signal-to-noise ratio exceeds 4 dB. Data records are transferred from the D-6000 data acquisition system to an HP-85 controller, and from the controller to magnetic tape cartridge mass storage device. Transfer from the data acquisition system to the mass storage tape drive requires about 2 seconds during which data acquisition is disabled. The system cycles through 4 different frequencies every 2 hours with a five minute noise measurement at the beginning of each 30-minute frequency period.

Data cartridges are returned for processing and analysis, the sequence for which is shown in Figure 1. Data from the tape cartridges consisting of raw voltage measurements are calibrated to received signal power and transferred to disc. A 20 item header consisting of the date, time, noise level, transmit power, frequency and other pertinent information is attached to each 512-point data record.

The next procedure involves identifying the dominant propagation mechanism in each record and if the dominant mechanism in the data record is meteor propagation, identifying the type (either underdense or overdense) of each meteor trail within the record. Several different propagation mechanisms are observed on the High Latitude Test Bed. In addition to underdense and over-
dense meteor trails, which due to differences in propagation mechanisms have different communication characteristics, sporadic-E and other low level ionospheric propagation are often observed. The test bed is well above the auroral and so auroral propagation is not observed.

After classification, data records are processed using the High Latitude Meteor Burst Data Analysis Package which creates a data base from the individual data records [Weitzen, 1986]. Information from the data base can be accessed to determine propagation and communication parameters of interest.

The classification procedure which is currently performed manually is the "weak link" in the data analysis procedure. It is an important function since the different propagation mechanisms and different types of meteor trails have different communication properties. In the classification procedure, each record is examined and the dominant propagation mechanism (either meteoric or ionospheric) is determined. If the dominant mechanism is meteor propagation, the type and time within the record of each meteor trail (either underdense or overdense) is identified. If there are no trails due to a false trigger, the data record is marked for discard.

Manual classification is both tedious and time-consuming, since between 15,000 and 60,000 data records per month are collected. The time required to classify the data from one month (60,000 records at 3 records per minute) would require approximately 330 work hours or about 2 months working 8 hours per day. Further, classification of events is often prone to human error, can vary from operator to operator, and can change with operator fatigue.
SECTION 2
AN AUTOMATIC CLASSIFICATION PROCEDURE

In order to automate the classification procedure, two tasks must be accomplished. First, the propagation mechanisms that are to be classified must be defined in terms of clearly identifiable characteristics. Second, a set of heuristics must be developed which emulate the processes of a human classifier.

2.1 PROPAGATION MODEL DEFINITION

The first step in developing an automatic classification program is to clearly define the characteristics of the propagation mechanisms that the program (or a human) is trying to identify.

The most frequently occurring events on meteor burst links are underdense meteor trails. Underdense meteor trails are characterized by their short rise time as the meteor passes through the zone of constructive interference followed by an exponential decay as the electrons in the trail diffuse [Manning, 1954]. An underdense meteor trail is modeled as a cylindrical Gaussian cloud of electrons in which each electron reflects independently of all others. Eshlemann [1955] developed accurate and relatively simple closed form solutions for the decay of underdense trails. Figures 2, 3 and 4 show examples of underdense trails at 45 MHz. Figure 5 shows a typical underdense trail at 104 MHz. In all the figures, the solid line represents the actual data and dashed lines represent attempted exponential fits to the data used to identify the type of meteor trail.

The key parameters of underdense trails, rise time, constant of decay and peak amplitude vary with known factors such as link distance, link power factor (transmit power antenna gains
etc.) and operating frequency and with factors which are not known a priori including trail orientation within the common volume, size of the meteor that formed the trail, and trail height and location within the common volume.

Overdense meteor trails occur less frequently but due to their high reflectivity and long duration, are potentially valuable for high throughput communication. Due to the increased electron density, the propagation model is much different and more complicated than that of underdense trails in which each electron is assumed to reflect independently of all others. Overdense trails are modeled to a first-order approximation as a cylindrical metallic tube [Hines and Forsythe, 1957]; however, this approximation has been shown to be relatively accurate only for the densest overdense trails. For meteor trails which are neither clearly underdense or so dense that the overdense approximation applies, there is no accurate and simple closed form equation for the decay. Overdense trails which last for one second or more usually fade due to a variety of mechanisms including multiple meteor trails and wind-induced distortion of the trail [Weitzen, 1984; Manning, 1959]. Examples of overdense trails are shown in Figures 6, 7 and 8. Overdense trails with fading are shown in Figures 9 and 19.

The difference between overdense and underdense meteor trails is of interest to communication engineers because the ground illumination footprint for overdense meteor trails is much larger than underdense trails reducing the inherent AJ (anti-jam) and LPI (low probability of intercept) in the channel. Due to their high reflectivity and long duration, overdense trails are potentially very valuable for high throughput communication. Underdense trails, due to their frequent occurrence, are valuable for rapid relay of short messages.
In addition to the classic underdense and overdense trails, several other types of meteoric echoes have been observed. On some very short duration meteor trails, one end of the meteor trail may decay before the other end has formed. The echo is due primarily to the region in the vicinity of the meteor head. The waveform rises and decays rapidly in less than 200 to 300 ms. These waveforms have been called "tiny" by [Ostergaard et al., 1985] and head echo trails by others [Eshlemann, 1960]. Figures 12 and 13 show examples of the very small meteors.

Large meteor trails, not oriented at formation so as to produce a specular reflection, can warp and drift to produce a reflection. These trails do not show the relatively sharp rise times associated with either underdense or classical overdense meteor trails and are often confused with ionospheric propagation. These echoes have been given the name "non-specular overdense trails" [Oetting, 1980]. Figures 10 and 11 show examples of non-specular overdense trails.

In the classification procedure, only the two classifications, underdense and overdense, are considered. Non-specular overdense trails are considered overdense and "tiny" meteors are considered underdense.

In addition to meteor propagation, at high latitudes ionospheric propagation and Sporadic-E propagation are commonly occurring phenomenon. Sporadic-E propagation is a relatively long-lasting propagation event characterized by slow fading and relatively high signal levels. Events can last from several minutes to hours or more. Lower level sporadic-E or ionospheric propagation is characterized by a continuous background with more rapid fading due to interference caused by small meteors. Ionospheric propagation has been observed at the 45 MHz frequency and has not been observed at the higher frequencies. Scatter from the auroral oval (located far to the south of the transmitter at
Sondrestrom) is not observed on this link. Figures 14, 15 and 16 show examples of low level and high level ionospheric propagation.
SECTION 3
THE AUTOMATIC CLASSIFICATION HEURISTICS

In the previous section we have defined the properties of the propagation mechanisms that are to be classified. In this section, the heuristics which make use of the previous definitions are described. The heuristic algorithm attempts to emulate the thought processes of a human classifier. It is implemented as a series of FORTRAN routines on a PDP-11/70 and a VAX/750. In the development of the algorithm, a number of thresholds and limits were set arbitrarily based on manual observation of several thousand data records. The levels were then adjusted to minimize the differences between the decision of the program and that of human classifiers.

The classification algorithm makes two passes through the data. On the first pass, each data record is classified as to the dominant propagation mechanism (meteoric, ionospheric, noise measurement, or false trigger) based only on information in the record and in previous records. In each record determined to be meteoric, all meteor trails within the record are classified as underdense or overdense.

In the second pass through the data, the consistency of the classification, based both on the classification of future and past records is checked. Changes in the classification are made when inconsistencies are detected.

3.1 IDENTIFYING THE DOMINANT PROPAGATION MECHANISM

The first step in the classification procedure involves identifying the dominant propagation mechanism in the data record. While the goal of the procedure is correct identification of the propagation mechanism, certain errors in classifi-
cation are more serious than others. In this procedure we define a series of "costs" associated with correct and incorrect identification. We will define the missed identification of ionospheric propagation $C(\text{meteoric} | \text{ionospheric})$ to be more severe than false identification of ionospheric propagation $C(\text{ionospheric} | \text{meteoric})$.

The high cost attributed to the error $C(\text{meteoric} | \text{ionospheric})$ is due to the fact that meteor burst propagation is a low duty cycle, intermittent event. A few extra seconds of high signal level ionospheric propagation mistakenly identified as meteoric could adversely affect the meteor burst communication statistics. $C(\text{ionospheric} | \text{meteoric})$ is less severe since ionospheric propagation tends to be a long term phenomenon and a few extra seconds would not adversely affect the communication statistics.

On the first pass through the data, the propagation mode identification algorithm has been designed to minimize the error condition of missed ionospheric propagation, $C(\text{meteoric} | \text{ionospheric})$. False classifications of ionospheric propagation caused by the desire to minimize missed ionospheric identifications are corrected on the second pass through the data.

The second pass through the data reduces the instances of incorrect classification by identifying inconsistencies in the classification. It uses the fact that ionospheric propagation tends to be a long duration event and the classification of both future and past records can be used to identify time-isolated ionospheric propagation records. These records are most likely non-specular overdense trails and are reclassified as meteoric and processed by the trail classifier. The second consistency check performed by pass 2 attempts to further reduce the number of missed ionospheric records by checking for records which precede the beginning of ionospheric events and may have not been identified by pass 1.
In the propagation identification phase of pass 1, once the beginning of an ionospheric event is identified, subsequent records can be classified as ionospheric with a high probability of being correct if they are time-adjacent. Identifying the beginning of an ionospheric event is difficult; however, several characteristics indicate ionospheric propagation with a fairly high probability.

First, if the envelope is greater than 10 dB above the noise level for more than 90% of the data record, it is probably ionospheric propagation and the record is given a preliminary identification of ionospheric propagation. Figure 15 shows such an example. If there are more than 6 fades in a four-second record and the average signal level in the record exceeds 4 dB above the noise, a preliminary identification of ionospheric propagation is made. Figure 14 shows an example of this phenomenon. If the average signal level in the record exceeds 4 dB above the noise but less than 8 fades are detected, then one final test is applied. In ionospheric propagation including fading, the temporal width of a fade is generally less than one second. If the width of a fade is greater than a second, then the record probably contains several large meteor trails. Figure 4 shows an example of a meteor trail that might be confused for ionospheric propagation were it not for the one-second criteria. Once the start of an ionospheric event is indicated, subsequent records with a peak signal level 4 dB above the noise are considered to be ionospheric if they are time-adjacent to a previous ionospheric record.

Records which are falsely identified as ionospheric which are non-specular overdense trails are corrected on the second pass. Records which precede the identified start of an ionospheric propagation event but which have not been classified as ionospheric on pass 1 are reclassified on pass 2. Records which are not ionospheric are given a default classification meteoric
and further processed in pass 1. Records corrupted by noise spikes or tape read or write errors are thrown out during the propagation mode identification.

Using this algorithm there was a difference of opinion between a committee of "expert" classifiers and the decision of the program (on fewer than 4% of the data records). The vast majority of the errors were of the less serious false ionospheric identification.

3.2 METEOR TRAIL CLASSIFICATION

After preliminary identification of the propagation mechanism in pass 1, records classified as meteoric are next processed to classify each meteor trail within the record. The algorithm for identifying the type of meteoric echo operates on the premise that if a trail is not specifically underdense (including tiny), the trail is overdense. This is the procedure used by human classifiers.

The first step in the meteor trail classification process requires the identification of the beginning and end of potential meteor trails. Potential meteor trails begin whenever the signal plus noise exceeds the minimum of 4 dB above the noise level (the approximate trigger level of the data acquisition system), and end when the signal level goes below the threshold. The time of the start of the trail, the end of the trail and the peak value are computed for each meteor trail. To prevent short noise bursts from being considered a meteor trail, potential trails must remain above the signal threshold for a prespecified amount of time. This time is a function of frequency (longer for lower frequencies) and was determined empirically. If no potential trails are identified in the preliminary pass the data record is marked for discard.
In the next pass, fades on meteor trails which may have been confused by the potential trail identifier as multiple trails are merged. If the duration between the end of an event and the beginning of the next event is less than about 150 ms, the cause is with a high probability fading rather than multiple trails and the two events are merged. On rare occasions (on 1% or less of the 2000 records observed) two very closely spaced trails are confused with fading (Figure 17) and on about 1% of the trails, very slow fading is confused as multiple trails. Figure 18 shows correct identification and classification of 3 meteor trails. Figure 19 shows the correct merging of fades on an a trail.

Having identified meteor trails and merged them to remove the effect of fading, the final task of the trail identifier is to decide whether each of the trails is underdense or overdense.

In the trail classification procedure, a trail is considered to be underdense or "tiny" if it satisfies a set of criteria and is considered overdense if it fails to satisfy the criteria. This premise is driven by the fact that closed form solutions for underdense trails exist while no closed form equations for overdense trails exist. Underdense meteor trails are characterized by their "short" rise time and exponential decay. Tiny meteors are characterized as having a total duration less than 200 ms. To test for the exponential decay, a least-square exponential curve fit is applied to each meteor trail from its peak value to the point where it is 3 dB above the noise floor. Relative error between the meteor trail and the curve fit is then computed. In addition to a "good" (less than 1 dB average error) fit to an exponential decay, a meteor echo must have a rise time to its max which is less than 250 ms or 25% of the total duration to be called underdense. Figure 2, 3, 4 and 5 show good fits and short rise times.
Any trail which has total duration less than 250 ms is automatically underdense ("tiny"). Figures 11 and 12 show trails which meet the short duration requirement. The rise time to max threshold is a function of frequency while the 25% criterion holds for all frequencies. Any trail which fails to meet either the "tiny" or the classic underdense criterion is identified as overdense. Figures 6 through 11 show examples of trails which have been classified as overdense. These criteria are applied to each trail in the four-second data window. If no trails are detected during the trail classification procedure the data record is marked for discard.

The trail classifier can be invoked by the pass two processor when the inconsistency of an isolated ionospheric propagation record is detected. The record is reclassified as meteoric and the trail classifier is invoked.

Figures 2 through 9 used previously to show examples of the various types of trails and events, show operation of the classifier. In the figures, the solid line represents the actual data in dBm and the dashed line shows the attempted exponential fit to each of the meteor trails. A third constant value solid line on the figures represents the average signal value described earlier.

3.3 EVALUATING THE ACCURACY OF THE CLASSIFIER

Determining the accuracy of the classification program is a difficult procedure because absolute knowledge of the propagation mechanism is not known. The next best procedure is to compare the classifications of the program to that of a group of "expert" classifiers.

As a first test, a group of five experienced operators manually classified 600 data records. The classification of each data record was determined by majority vote of the group. A typ-
ing error was declared if the autoclassification program disagreed in the identification of the dominant propagation mode and a trail error was declared if the autoclassifier failed to identify or incorrectly identified the type of a meteor trail. On records in which the group could not agree on a classification (defined as 3 to 2 vote), the record or trail was removed from the statistics.

The automatic classifier disagreed with dominant propagation mode classification of the committee on 24 (4%) of the 600 records and disagreed with the meteor trail identification of the committee on 39 out of 780 (5%) meteor trails. The primary errors in propagation mechanism identification were records falsely typed as ionospheric which were actually noise bursts or rapidly fading meteor trails. The errors were not identified on the second pass due to the fact that meteoric records may have been time-adjacent. This error C(ionospheric|meteoric or throw out) is defined to not be a serious error. In only 1 out of 600 records was an ionospheric event missed C(meteoric|ionospheric).

The primary error in typing meteor trails occurred when trails faded so slowly that they were falsely counted as separate trails. The other type of error which occurred was underdense trails classified as overdense because fading caused them to have a "poor" fit to an exponential.

A second performance test involved comparing the autoclassifier to a single relatively inexperienced classifier. Before the autoclassifier, various operators would classify data records in eight hour shifts. A computer program was written to compare the results of the manual classifier and the autoclassifier. The autoclassifier disagreed with the dominant propagation mode identification of the operator about 8% of the time. In about two-thirds of the differences, the committee observed that the human classifier had erred. The autoclassifier disagreed with the manual classifier on about 12% of the trail
identifications. Once again the committee observed that at least 1/2 of the differences were due to operator error.

In order to further improve the accuracy of the classification procedure, RADC/EEPS personnel make a manual pass through the data. The consistency of classification changes made on pass two through the data is examined manually. In addition, all records in which three meteor trails are classified are reviewed manually. This procedure requires manual classification of about 4% of the data records and reduces the classification errors to less than 2%. The third manual pass through the data will be eventually phased out.
SECTION 4
CONCLUSIONS

Automatic classification of data records for the Air Force High Latitude Meteor Burst Test Bed, while not accurate 100% of the time, allows timely analysis of large quantities of data from high latitude meteor burst test bed. In comparisons with "expert classifiers" the program disagrees with the committee on about 5% of the records. In comparison to a less experienced operator operating for long periods of time, the program performance was superior. The program reduces the time required to process one month of data from 2 staff months when classified manually to about 12 hours when classified automatically.
SECTION 5
ACKNOWLEDGMENT

The authors wish to acknowledge the members of the "expert" classification committee at RADC/EEPS John Rassmussen, Dr. Paul Kossey, Mike Sowa, Jake Quinn, Wayne Klemente, Lt. Rob Scofidio, and SSGT Wade Warrens for their help and advise during the development process. The work was performed under U.S. Air Force Contract F19628-84-C0117.
REFERENCES


Weitzen, Jay A., A data base approach to analysis of meteor burst data. To be published.
Figure 1 Flow diagram of high latitude meteor burst test bed data reduction and analysis.
Figure 2  Underdense meteor trail at 45 MHz.
Figure 3  Two underdense meteor trails at 45 MHz.
Figure 4. Very large underdense meteor trail at 45 MHz.
Figure 5  Underdense meteor trail at 104 MHz.
Figure 6  Overdense meteor trail at 45 MHz.
Figure 7. Overdense meteor trail at 45 MHz.
Figure 8  Overdense meteor trail at 65 MHz.
Figure 12 "Tiny" meteor trail at 45 MHz.
Figure 14 Example of low level ionospheric propagation event.
Figure 15 Example of high level ionospheric propagation.
Figure 17 Example of two closely occurring trails which are falsely merged into one trail.
Figure 18 Example of three correctly identified meteor trails in one four second window. (L to R underdense, overdense, underdense).
Figure 19 Correct merging of fades on fading meteor trail.