OAK RIDGE NATIONAL LABORATORY LIBRARIES

3 4456 0536019 6

ORNL/TM-2002/183

OAK RIDGE
NATIONAL LABORATORY
MANAGED BY UT-BATTELLE
FOR THE DEPARTMENT OF ENERGY

# Annual Report

## Nuclear Energy Research and Development Program
## Nuclear Energy Research Initiative

## NERI Proposal #2000-0109
## Forewarning of Failure in Critical Equipment at Next-generation NPP

## August 2001–2002

L. M. Hively
V. A. Protopopescu

December 2002

UT-BATTELLE
ORNL-27 (4-00)

DOCUMENT AVAILABILITY

Reports produced after January 1,1996, are generally available free via the U.S. Department of Energy (DOE) Information Bridge.

*Web site* http:/www.osti.gov/bridge

Reports produced before January 1, 1996, may be purchased by members of the public from the following source.

National Technical Information Service
5825 Port Royal Road Springfield, VA 22161
*Telephone* 703-605-6(XX) (1-800-553-6847)
*TDD* 703-487-4639
*Fax* 703-605-6900
*E-mail* info@ntis.fedworld.gov
*Web site* http:/www.ntis.gov/support/ordernowabout.htm

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE) representatives, and International Nuclear Information System (INIS) representatives from the following source.

Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831
*Telephone* 865-576-8401
*Fax* 865-576-5728
*E-mail* reports@adonis.osti.gov
*Web site* http:/www.osti.gov/contact.html

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/TM-2002/183

**Computational Sciences and Engineering Division**

**Annual Report**

**Nuclear Energy Research and Development Program**
**Nuclear Energy Research Initiative**

**NERI Proposal #2000-0109**
**Interoffice Work Order (IWO) M0SF 00-109**
**Forewarning of Failure in Critical Equipment at Next-Generation NPP**

**August 2001–August 2002**

L.M. Hively
V.A. Protopopescu

December 2002

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6285
managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
Under contract DE-AC05-00OR22725

## CONTENTS

Page

ABSTRACT ..... viii

LIST OF TABLES ..... iv

LIST OF FIGURES ..... v

1. ADDITIONAL DATA ..... 1
   1.1 ANALYSIS OF EPRI MOTOR POWER DATA: AIRGAP-OFFSET FAULT ..... 3
   1.2 ANALYSIS OF EPRI MOTOR POWER DATA: BROKEN-ROTOR FAULT ..... 5
   1.3 ANALYSIS OF EPRI MOTOR POWER DATA: TURN-TO-TURN SHORTS ..... 6
   1.4 ANALYSIS OF DATA FROM POLYTECHNIC UNIVERSITY OF VALENCIA ..... 7
   1.5 ANALYSIS OF MOTOR DATA FROM PSU ..... 8
   1.6 ANALYSIS OF MOTOR DATA FROM PSU (RUN 33) ..... 9
   1.7 ANALYSIS OF MOTOR DATA FROM PSU (RUN 34) ..... 11
   1.8 ANALYSIS OF MOTOR DTAT FROM PSU RUN 35) ..... 12
   1.9 ANALYSIS OF TORSION DATA FROM PSU ..... 12
   1.10 ANALYSIS OF BEARING DATA FROM PSU ..... 13

2. DISCUSSION ..... 16
   2.1 LESSONS LEARNED ..... 16
   2.2 OTHER WORK ..... 17
   2.3 ISSUES/CONCERNS ..... 19
   2.4 COST PERFORMANCE ..... 20

APPENDIX A: DESCRIPTION OF ANALYSIS METHODS ..... 66
APPENDIX B: MDTB TEST PLAN ..... 76
APPENDIX C: TEST PLAN FOR TORSION EXPERIMENT ..... 83
APPENDIX D: BEARING PROGNOSTIC TEST RIG ..... 90

## LIST OF TABLES

**Table** **Page**

1. Summary of Contacts for Additional Test Data ..... 2
2. First Motor Specifications ..... 3
3. General Electric Motor Specifications ..... 6
4. Data and Failures ..... 16
5. Further Details of new Data Acquisition System ..... 18
6. Status Summary of NERI Tasks for First and Second Project Years ..... 21

## LIST OF FIGURES

**Figure** **Page**


1. Motor-bearing system (top), bearing details (bottom left), and bear assembly (bottom right) ..... 22
2. Baseline data vs time from the equipment in Fig. 1 ..... 23
3. Two dimensional phase-space plots for acoustic data in top plot of Fig. 2 ..... 24
4. Typical baseline data vs time from the Allis Chalmers motor ..... 25
5. Linear measures for the air-gap seeded-fault ..... 26
6. Conventional nonlinear measures for the air-gap seeded-fault ..... 27
7. Two-dimensional phase-space reconstructions for baseline power data ..... 28
8. PSDM for the airgap-offset seeded-fault ..... 29
9. Linear measures for the broken-rotor seeded-fault ..... 30
10. Conventional nonlinear measures for the broken-rotor seeded-fault ..... 31
11. PSDM for the broken-rotor seeded ..... 32
12. Typical baseline data vs time from the GE motor ..... 33
13. Linear measures for the turn-to-turn short seeded-fault ..... 34
14. Conventional nonlinear measures for the turn-to-turn short seeded-fault ..... 35
15. PSDM for the turn-to-turn short seeded-fault ..... 36
16. Conventional statistical measures vs dataset number ..... 37
17. Traditional nonlinear measures vs time ..... 38
18. Unrenormalized PSDM vs dataset number ..... 39
19. Mechanical Diagnostics Test Bed at PSU/ARL ..... 40
20. Lnear measures of load torque from PSU/ARL MDTB ..... 41
21. Cnventional nonlinear measures of load torque data ..... 42
22. PSDM vs dataset number ..... 43
23. Raw tri-axial accelerometer data ..... 44
24. Conventional statistical measures of tri-axial acceleration ..... 45
25. Tri-axial accelerometer power vs time ..... 46
26. Conventional statistical measures of the accelerometer power ..... 47
27. Conventional nonlinear measures vs time ..... 48
28. PSDM vs dataset number ..... 49
29. Maximum value $\chi^2$ vs number ($n$) of sequential points ..... 50
30. Tri-axial accelerometer power vs time ..... 51
31. Conventional statistical measures of accelerometer power vs dataset number ..... 52
32. Conventional nonlinear measures vs time ..... 53
33. PSDM vs dataset number ..... 54
34. Maximum value $\chi^2$ vs number ($n$) of sequential points ..... 55
35. Four data channels from PSU TORSION data ..... 56
36. Conventional statistical measures of DC motor power vs dataset number ..... 57
37. Conventional nonlinear measures vs time ..... 58
38. PSDM (top four) and their sum (bottom) vs dataset number ..... 59
39. Accelerometer power from the PSU BEARING experiment ..... 60
40. Conventional statistical measures of accelerometer power vs dataset number ..... 61
41. Conventional nonlinear measures vs time ..... 62
42. PSDM (top four) and their sum (bottom) vs dataset ..... 63

## LIST OF FIGURES (continued)

**Figure** **Page**

43. Maximum $\chi^2$ vs the number ($n$) of sequential points ..... 64
44. Budget and schedule ..... 65

## ABSTRACT

NERI Project #2000-0109 began in August 2000 and has three tasks. The first project year addressed Task 1, namely development of nonlinear prognostication for critical equipment in nuclear power facilities. That work is described in the first year's annual report (ORNL/TM-2001/195). The current (second) project year (FY02) addresses Task 2, while the third project year will address Tasks 2–3. This report describes the work for the second project year, spanning August 2001 through August 2002, including status of the tasks, issues and concerns, cost performance, and status summary of tasks.

The objective of the second project year's work is a compelling demonstration of the nonlinear prognostication algorithm using much more data. The guidance from Dr. Madeline Feltus (DOE/NE-20) is that it would be preferable to show forewarning of failure for different kinds of nuclear-grade equipment, as opposed to many different failure modes from one piece of equipment. Long-term monitoring of operational utility equipment is possible in principle, but is not practically feasible for the following reason. Time and funding constraints for this project do not allow us to monitor the many machines (thousands) that will be necessary to obtain even a few failure sequences, due to low failure rates ($<10^{-3}$/year) in the operational environment. Moreover, the ONLY way to guarantee a controlled failure sequence is to seed progressively larger faults in the equipment or to overload the equipment for accelerated tests. Both of these approaches are infeasible for operational utility machinery, but are straight-forward in a test environment. Our subcontractor has provided such test sequences. Thus, we have revised Tasks 2.1–2.4 to analyze archival test data from such tests.

The second phase of our work involves validation of the nonlinear prognostication over the second and third years of the proposed work. Recognizing the inherent limitations outlined in the previous paragraph, Dr. Feltus urged Oak Ridge National Laboratory (ORNL) to contact other researchers for additional data from other test equipment. Consequently, we have revised the work plan for Tasks 2.1–2.2, with corresponding changes to the work plan as shown in the Status Summary of NERI Tasks (below). The revised tasks are as follows.

Task 2.1: ORNL will obtain test data from a subcontractor and other researchers for various test equipment. This task includes development of a test plan or a description of the historical testing, as appropriate: test facility, equipment to be tested, choice of failure mode(s), testing protocol, data acquisition equipment, and resulting data from the test sequence. ORNL will analyze this data for quality, and subsequently via the nonlinear paradigm for prognostication.

Task 2.2: ORNL will evaluate the prognostication capability of the nonlinear paradigm. The comparison metrics for reliability of the predictions will include the true positives, true negatives, and the forewarning times.

Task 2.3: ORNL will improve the nonlinear paradigm as appropriate, in accord with the results of Tasks 2.1–2.2, to maximize the rate of true positive and true negative indications of failure. Maximal forewarning time is also highly desirable.

Task 2.4: ORNL will develop advanced algorithms for the phase-space distribution function (PS-DF) pattern change recognition, based on the results of Task 2.3. This implementation will provide a capability for automated prognostication, as part of the maintenance decision-making.

Appendix A provides a detailed description of the analysis methods, which include conventional statistics, traditional nonlinear measures, and ORNL's patented nonlinear PSDM. The body of this report focuses on results of this analysis.

# 1. ADDITIONAL DATA

Dr. Feltus suggested potential contacts for additional data. Oak Ridge National Laboratory (ORNL) has pursued these and follow-on contacts as summarized in Table 1. To date, ORNL has obtained data from several researchers, as discussed next.

Don Jarrell of Pacific Northwest National Laboratory (PNNL) sent information to ORNL via e-mail on 10/2/01, including a description of pump cavitation tests from their NERI project. The specific data spanned 600–1400 s for each of three experiments. However, the sampling rate of 1 Hz was much too slow to allow analysis by our nonlinear algorithm. No additional data has been received from Don Jarrell.

ORNL obtained test data on CD-ROM from Dr. Amir Shirkhodaie (Tennessee State University, Department of Mechanical Engineering) on November 17, 2001. These tests involve a motor-bearing system (Fig. 1) running at 900–3500 rpm. Sixteen channels of data were recorded in each record, including a time stamp, twelve accelerations, two forces, and one acoustic reading. All data were sampled at 10.8 kHz. Data quality analysis revealed one test sequence with only eight undocumented channels. Other data had blank records, inadequate dataset lengths, or an inconsistent number of channels across multiple datasets for the same test sequence. The data quality check was passed by one test sequence for which typical data segments are shown in Fig. 2. Only a single (acoustic) channel in this sequence see med to have an adequate sampling rate (top plot in Fig. 2). This test involved an imbalance fault (130 grams) at 1500 rpm with data sampled at 10.8 kHz for 5 s (54,000 data points). Two such imbalance datasets were provided, along with two normal datasets, one before the imbalance tests and one afterward. However, the two baseline datasets have different features, just by comparing the linear measures: minimum, maximum, absolute average deviation, standard deviation, skewness, kurtosis, average number of time steps per cycle, and first zero in the auto-correlation function. Unsurprisingly then, the nonlinear measures of dissimilarity between the baseline datasets are large (>0.7 standard deviations from the mean) compared to the intra-dataset variability in those same nonlinear dissimilarity measures (≤0.5 standard deviations from the mean). Comparison of both test cases to the second baseline shows clear differences in the linear measures by a factor of 2–2.5. Nonlinear measures of dissimilarity between the second baseline and the test cases are 3–5 standard deviations from the mean, which is better than the linear measures. Figure 3 shows two-dimensional phase-space portraits of this acoustic data, $x_i$, by plotting pairs of points ($x_{i+\lambda}$, $x_i$) that are connected by straight lines for values of lag, $1 \leq \lambda \leq 20$. The images resemble a cross, indicating abrupt changes due to an inadequate sampling. ORNL sent an e-mail to Dr. Shirkhodaie on November 12, 2001 with details of this analysis, questions, and a suggestion for increasing the data sampling rate. Dr. Shirkhodaie's response indicated this data was from initial experiments, and that they are working to eliminate the data quality problems. Dr. Shirkhodaie said in a subsequent e-mail exchange that he would be glad to share additional data with ORNL as it becomes available. No additional data has been received from Dr. Shirkhodaie.

A third fruitful contact is Jan Stein (EPRI) who was the project leader for evaluation of various commercial diagnostics for large motors, beginning in 1994. The project report is "Electric Motor Predictive Maintenance Program," TR-108773-V2 (1999), which ORNL bought from

## Table 1. Summary of Contacts for Additional Test Data

| Contact's Name | e-mail address | Phone number | Suggested by | Outcome |
|---|---|---|---|---|
| Leonard Bond (PNL) | Leonard.Bond@pnl.gov | 509-375-4486 | Madeline Feltus (DOE) | Check Don Jarreli |
| Belle Upadhyaya (UTK) | bupadhya@utk.edu | 865-974-5048 | Madeline Feltus (DOE) | Check Tom Byerly (UTK) |
| Frank Rahn (EPRI) | frahn@epri.com | 650-855-2037 | Madeline Feltus (DOE) | Provided other contacts |
| Lance Agee (EPRI) | lagee@epri.com | 650-855-2106 | Madeline Feltus (DOE) | Check Frank Rahn |
| Nathan Siu (NRC) | nos@nrc.gov | 301-415-6952 | Madeline Feltus (DOE) | Provided other contacts |
| Steven Arndt (NRC) | saa@nrc.gov | 301-415-6502 | Madeline Feltus (DOE) | Provided other contacts |
| Bob Uhrig (UTK) | ruhrig@utk.edu | | Steven Arndt (NRC) | Check Wes Hines |
| Prof. Moran (Ohio State) | | 614-292-6064 | Steven Arndt (NRC) | Provided other contacts |
| Prof. Jack Collins (Ohio State) | | 614-292-6094 | Prof. Moran (Ohio State) | Unable to participate |
| Kent Hansen (MIT) | | 617-253-7384 | Nathan Siu (NRC) | Provided other contacts |
| George Apostolakis (MIT) | apostola@mit.edu | 617-252-1570 | Kent Hansen (MIT) | no suggestions |
| Jim Lynch (INPO) | | 770-644-8000 | Kent Hansen (MIT) | No response |
| Ken Barry (EPRI) | | 704-547-6040 | Frank Rahn (EPRI) | Check Frank Rahn |
| Ken Huffman (EPRI) | | 704-547-6055 | Frank Rahn (EPRI) | No response |
| Richard Wood (ORNL) | w15@ornl.gov | 865-574-5578 | Leonard Bond (PNL) | Check Bel Upadhyaya (UTK) |
| Leonard Loflin (EPRI) | leloflin@epri.com | 704-547-6010 | Ken Barry (EPRI) | Interested |
| Tom Byerly (UTK/MCR) | | 865-974-9625 | Belle Upadhyaya (UTK) | Presentation to MRC at UTK |
| Wes Hines (UTK) | hines@utkux.utcc.utk.edu | 865-974-6561 | Bob Uhrig (UTK) | Sample rate too low |
| Darryl Cox (ORNL) | | | Ray Holdaway (ORNL) | No response |
| Al Wilks (Smart Signal) | adwilks@smartsignal.com | | Presentation to MRC | Interested |
| Bill Drake (Wilcoxon Res) | bdrake@wilcoxon.com | 301-216-3020 | Presentation to MRC | Interested |
| Bill Adams (Fuller) | | | Presentation to MRC | Interested |
| Sam Robinson (Y-12 Plant) | | 865-574-1838 | Previous collaboration | Check Al Akerman |
| Roger Kershaw (MachinExpert) | | 865-637-1760 | Cecil Presnell (IdentiChem) | Interested |
| Ken Piety (CSI) | | 865-675-2110 | Ray Garvey (CSI) | Only 4096-point datasets |
| Amulya Garga (PSU/ARL) | | 814-863-5841 | Amir Shirkhodaie (Tn St) | Check Karl reichard |
| Tim Rangongo (Iris Power) | | 416-620-5600 | Madeline Feltus (DOE) | Check Foster-Miller Tech. |
| Gordon Hirschman (FMT) | fhirschman@fosmiltech.com | | Tim Rangongo (Iris Power) | No response |
| Jan Stein (EPRI) | jstein@epri.com | 650-855-2390 | Madeline Feltus (DOE) | EPRI provided data on CD |
| Don Jarrell (PNL) | Don.jarrell@pnl.gov | 509-372-4096 | Leonard Bond (PNL) | Promised to provide data |
| Mark Linn (HFIR/ORNL) | linnma@ornl.gov | 865-574-4617 | Dwayne Kilpatrick (ORNL) | Provided motor specs |
| Amir Shirkhodaie (TSU) | ashirkhodaie@tnstate.edu | 615-963-5396 | MARCON2001 proceedings | Provided data |
| Karl Reichard (PSU) | kmr5@psu.edu | 814-863-7681 | Amir Shirkhodaie (TSU) | Providing data |
| Al Akerman (ORNL) | ima@ornl.gov | 865-406-6589 | Sam Robinson (Y-12 plant) | Provided data |
| Maria Elena Otero (Spain) | memontes@iqn.upv.es | | Raphael Peres (ORNL) | Provided data |
| Michael House (GE) | Michael.house@ps.ge.com | 518-385-4098 | Jan Stein (EPRI) | Willing to share data |

EPRI for $1,000. The report included a CD-ROM of actual data from the testing. We expect that additional data will be forthcoming. ORNL's analysis of the present EPRI data is provided below.

We obtained sample data from Dr. Maria Elena Montesino Otero (Universidad Politeckica de Valencia, Spain). Our analysis of this data also is described below.

Dr. Al Akerman (ORNL) kindly provided machine data on CD-ROM. The data were acquired from vacuum pumps, blower motors, and purifiers in Building 9204-4 at the Oak Ridge Y-12 Plant. The specific data included motor current, vibration, moisture, oxygen, and gas flow rate. Most of the data were encoded in LabView™ binary files. RMS values of other data are at 1 min intervals, which is very under-sampled for our analysis. We are presently unable to pursue further analysis of this data.

## 1.1 ANALYSIS OF EPRI MOTOR POWER DATA: AIRGAP-OFFSET FAULT

EPRI began a project in 1994, entitled "Electric Motor Predictive Maintenance Program." Jan Stein (EPRI) led the project, which evaluated of various commercial diagnostics for large motors. Details are reported in EPRI report # TR-108773-V2 (1999). The report included a CD-ROM of actual data from the testing. That work involved collaboration by several utilities and EPRI on seeded faults in large electric motors. The datasets were recorded in snap-shots of 1.5 s, sampled at 40 kHz (60,000 total time-serial samples), including three-phase voltages and currents, plus tri-axial accelerations at inboard and outboard locations on the motor. Several anomalies were introduced in the motors to simulate the most common pre-failure in-service conditions. ORNL received data for three different seeded faults via the CD-ROM that accompanied the EPRI report. Table 2 shows the specifications of the first motor.

Table 2. First Motor Specifications

| | | | |
|---|---|---|---|
| Manufacturer: | Allis Chalmers | Bearing Type: | Sleeve |
| Rated voltage: | 4160 | Nameplate current: | 100 amps |
| Rated hp: | 800 | Number of rotor bars: | 94 |
| Winding type: | Form wound | Number of stator slots: | 94 |
| Phases: | 3 | Hertz: | 60 |
| Rpm: | 710 | Motor type: | Induction |
| Insulation class: | F | Poles: | 10 |
| Enclosure: | TEFC | Bar configuration; | copper |

The test sequence initially acquired data from the motor running in its nominal state (first dataset). Two different airgap-offset seeded-faults then were imposed via preinstalled jackscrews. The second dataset involved an inboard airgap offset of 8 mils from the nominal value of 30 mils. The third dataset retained the first fault, and added an additional seeded-fault outboard airgap offset by 20% in the opposite direction from the inboard shift. This additional fault

resulted in the rotor being skewed relative to the stator. These offsets were static because neither varied relative to the stator with the motor running.

We discuss details of the data next. All of the datasets on the EPRI CD-ROM had non-printable characters at the beginning and end of the file that confounded normal input as ASCII data to the FORTRAN analysis. Moreover, the 'TYPE' utility in DOS could not display the non-printable data. However, the MATLAB™ editor was able to read these datasets. This editor was also able to write a FORTRAN-readable ASCII file after adding, then deleting, a space at the beginning of the file, and after the addition of a carriage-return/line-feed at the end of the file. The three datasets for this first test sequence were concatenated into a single long dataset for ease of analysis. We converted three-phase voltages ($V_i$) and currents ($I_i$) into instantaneous motor power, $P = \Sigma_i I_i V_i$, where the sum runs over the three phases, as shown in Fig. 4. The particularly noteworthy feature in Fig. 4 is that the instantaneous power (bottom subplot) displays rich dynamical features, which are not at all apparent from the sinusoidally varying currents and voltages (upper six subplots) in Fig. 4. Consequently, we focus the analysis on instantaneous power, rather than the individual currents or voltages. Closer examination of the power reveals a slow, low-amplitude variation with a period of roughly 0.1 s. We removed this artifact with ORNL's patented zero-phase quadratic filter. Otherwise, this artifact confounds the interpretation of our results. We split each of the three datasets into five subsets of 12,000 points each, giving fifteen total subsets for analysis. This artifact-filtered data showed no data quality problems.

We next compare the condition change in linear, traditional nonlinear, and phase-space dissimilarity measures (PSDM). Figure 5 shows linear measures of artifact-filtered motor power. All of the linear measures show an abrupt change with the onset of the airgap-offset faults, but no other systematic trends. Figure 6 shows the change in three conventional nonlinear measures. The mutual information function (MIF) measures the average information (in bits) that can be inferred from one measurement about a second measurement and is a function of the time delay between the measurements. Univariate MIF measures predictability within the same data stream at different times. For the present analysis, we use the first minimum in the univariate MIF, $M_1$, to indicate the average time lag that makes $P_i$ independent of $P_j$. The top plot of Fig. 6 shows that $M_1$ decreases erratically from 13 to 11 time steps during the first (nominal) dataset, then varies erratically between 10 to 12 time steps during the second (8 mil air-gap test) dataset, and finally rises even more erratically from 12 to 16 time steps during the last (20% air-gap test) dataset. The middle plot of Fig. 6 shows the maximum likelihood correlation dimension ($D$), which is roughly constant at 3.7 during the first dataset, then decreases to a second roughly constant value at 1.4 for the second and third datasets. The bottom plot displays the maximum-likelihood Kolmogorov entropy ($K$), which measures the rate of nonlinear information loss (bits/s) in the data. K rises erratically from 0.019 to 0.026 during the first dataset, then decreases to roughly constant value of nearly zero for the second and third datasets. Figure 7 illustrates the two-dimensional phase-space reconstruction of the artifact-filtered nominal-state power for several different time lags, showing progressive "unfolding" of geometric representation for these dynamics.

A systematic search revealed a set of phase-space reconstruction parameters that are most sensitive to the condition change for the air-gap seeded-fault test sequence, as shown in Fig. 8.

This figure shows an almost linear rise in the connected-phase-space dissimilarities (second from the top and bottom plots) from near zero for the nominal state to approximately 20 for the double-seeded air-gap fault. The rise in non-connected dissimilarity measures is monotonic, but little changed between the two faulted states. Thus we find that condition change in this test sequence is detected to some degree by all of the linear, conventional nonlinear, and PSDM; the connected-phase-space dissimilarities show the progressively worsening severity of the fault most clearly.

## 1.2. ANALYSIS OF EPRI MOTOR POWER DATA: BROKEN-ROTOR FAULT

A second test sequence for this motor began with the same motor running in its nominal state (first dataset), followed by progressively more severe broken rotor bars. The second dataset involved a simulated failure that was one rotor bar cross section cut 50% in half at the 11 o'clock position. The third dataset was for the same rotor bar next cut through 100%. The fourth dataset was for a second rotor bar cut 100% at the 5 o'clock position, exactly 180° from and in addition to the first rotor failure. The fifth dataset was for two additional rotor bars cut adjacent to the original 11 o'clock bar, with one bar cut on each side of the original, yielding four bars completely open. The EPRI report says that the data-collection personnel noted a definite growling sound and a pulsating vibration during this last test. We concatenated the five datasets into a single long dataset for ease of analysis, and converted the three-phase voltages and currents into instantaneous power, as before. We split each of the five datasets into five subsets of 12,000 points each, giving 25 total subsets. The power has a slow, low-amplitude variation with a period of roughly 0.1 s. As before, we removed this artifact, which otherwise confounds the interpretation of our results. A check of this artifact-filtered data revealed no data quality problems.

We first show linear measures of artifact-filtered motor power in Fig. 9. The top plot in Fig. 9 illustrates that the minimum, maximum and standard deviation in motor power are essentially flat until the abrupt change during the last test (four cut rotor bars) in the sequence. Skewness and kurtosis (second plot down) also are flat until the last test, when they both change abruptly and in opposite directions. The number of time steps per cycle in the motor power (third plot down) decreases slowly and very noisily from 53 to 45 time steps per cycle (15% change). The first zero in the autocorrelation function varies erratically between 32–33 time steps during the nominal state, remains constant at 33 time steps for the first four faults, and then varies erratically again between 32–34 time steps during the last test.

For additional comparison, we show the variation of three conventional nonlinear measures in Fig. 10. The top plot of Fig. 10 shows that the first minimum in the mutual information function, $M_1$, varies erratically for the first four datasets, then rises to 14 time steps and remains there. The middle plot of Fig. 10 shows that the maximum likelihood correlation dimension ($D$) also varies erratically over the whole test sequence. The bottom plot displays the Kolmogorov entropy ($K$), which shows no consistent trend.

We systematically varied the phase-space reconstruction parameters to obtain the most monotonic increase in condition change for the broken-rotor seeded-fault test sequence.

Figure 11 shows that the phase-space dissimilarity measures rise by ten-fold over the test sequence. The parameters are: $S = 88$ (number of equiprobable phase-space symbols), $d = 4$ (number of phase-

space dimensions), $\lambda = 31$ (time delay lag in time steps), and $w = 550$ (half width of the artifact filter window in time steps). The exponential rise in the magnitude of the seeded faults (doubling from 0.5 to 1.0 to 2.0 to 4.0) is mirrored in Fig. 11 by a linear rise in the logarithm of the dissimilarity measures. We conclude that the conventional statistical and traditional nonlinear measures provide no indication of condition change due to the broken-rotor bars. In sharp contrast, PSDM show condition change that is a proportional to the exponentially increasing severity of the rotor-bar fault, providing clear indication of the failure.

## 1.3 ANALYSIS OF EPRI MOTOR POWER DATA: TURN-TO-TURN SHORTS

The EPRI data included a General Electric motor, as specified in Table 3.

**Table 3. General Electric Motor Specifications**

| | | | |
|---|---|---|---|
| Rated voltage: | 4000 | Bearing type: | Sleeve |
| Rated hp: | 500 | Number of rotor bars: | 84 |
| Winding type: | Form wound | Number of stator slots: | 108 |
| Phases: | 3 | Hertz: | 60 |
| Rpm: | 1185 | Motor type: | Induction |
| Insulation class; | B | Poles: | 6 |
| Enclosure: | Open | Bar configuration: | Copper rectangular |

This sequence began with the motor running in its nominal state (first dataset). The second dataset had a turn-to-turn (2.70 ohm) short by installing a large screw between two turns. The third dataset had a more severe turn-to-turn (1.35 ohm) short by installing a smaller screw between two turns. These three datasets were concatenated into a single long dataset for this analysis. The sequence for the concatenation goes from largest turn-to-turn resistance (infinite resistance, corresponding to no short), to smaller (2.7 ohms), to smallest (1.35 ohms), corresponding to increasing severity in the fault. As before, the three-phase voltages and currents were converted into instantaneous power. The three datasets were split into 5 subsets of 12,000 points each, giving 15 total subsets. Figure 12 shows three-phase voltages and currents, plus motor power for the nominal operating state. This data has a low-amplitude, low-frequency artifact with a period of roughly 0.006 s. As before, we remove this artifact, which would otherwise confound the interpretation of our analysis. The artifact-filtered power data has no data quality problems.

As before, our analysis compares the condition change in linear, traditional nonlinear, and PSDM. Figure 13 shows linear measures of artifact-filtered motor power. The minimum, maximum, and average (plus or minus one standard deviation) in the power (top subplot in Fig. 13) are essentially constant during the entire test sequence. The same is true of the skewness and kurtosis (second plot down in Fig. 13). The number of time steps per cycle varies erratically

around 17 throughout the test sequence. The first zero in the autocorrelation function also varies erratically between 11 and 28 during the first dataset, then is essentially constant during the fault tests.

Figure 14 shows the change in the three conventional nonlinear measures. The top plot of Fig. 14 shows that $M_1$ varies erratically between 8 and 10 time steps during the nominal state, decreases erratically to 7 time steps during the second dataset, and is roughly constant at 7–8 time steps during the last dataset. The middle plot of Fig. 14 shows the maximum likelihood correlation dimension ($D$), which varies erratically between 3.7 and 4 during the nominal state, varies erratically between 3.6 and 4.2 during the second dataset, and shifts from 4.08 down to a plateau of roughly 3.8 during the last dataset. The bottom plot displays the maximum-likelihood Kolmogorov entropy ($K$), which varies erratically between 0.043 and 0.079 during the first (nominal) dataset, varies between 0.06 and 0.1 during the second datasets, and is even more erratic (between 0.05 and 0.14) during the last dataset.

Systematic variation of the phase-space reconstruction parameters revealed values that are most sensitive to the condition change for the turn-to-turn seeded-fault test sequence. Figure 15 shows that all four of the PSDM rise monotonically over the test sequence. The parameters are: $S = 129$ (number of equiprobable phase-space symbols), $d = 3$ (number of phase-space dimensions), $\lambda = 1$ (time delay lag in time steps), and $w = 221$ (half width of the artifact filter window in time steps). The linear rise in the magnitude of the seeded faults (from 2.7–1.35 ohms) is mirrored in Fig. 15 by a linear rise in the dissimilarity measures. We conclude that the conventional statistics and traditional nonlinear measures provide no indication of failure for the turn-to-turn shorts. In sharp contrast, the nonlinear dissimilarity measures show condition change that is a proportional to the increasing severity of the rotor-bar fault, providing clear forewarning of the failure.

## 1.4 ANALYSIS OF DATA FROM POLYTECHNIC UNIVERSITY OF VALENCIA

Dr. Maria Elena Montesino Otero (Universidad Politeckica de Valencia, Spain) sent nine digital datasets via e-mail to ORNL on April 19, 2002. Each ASCII data file contained 16,384 points of uni-axial accelerometer data with corresponding time stamps. The data were sampled at 600 Hz from a three-phase, ¼-hp motor running at 18 Hz with a progressively larger imbalance fault. ORNL concatenated the nine datasets into one long data file for subsequent analysis. A check of the data quality revealed no data gaps (based on the time stamps) or other problems.

Figure 16 displays the conventional statistical measures for this data. Figure 16a shows a monotonic increase in the magnitude of the minimum, maximum, absolute average deviation, and the sample standard deviation. Figure 16b shows no correlation of the skewness with the dataset number. Figure 16c also shows a monotonic increase in the magnitude of the (negative) kurtosis with dataset number. The number of time steps per cycles also increases monotonically (Fig. 16d) over datasets 1–6, and then is flat for the remainder of the test sequence. The first zero in the auto-correlation function (Fig. 16e) is constant at eight time steps over the entire test sequence. Thus, some (but not all) of the conventional statistical measures are correlated with the increasing fault severity for this test sequence.

Figure 17 illustrates the three conventional nonlinear measures for this data. The top subplot in Fig. 17 shows that correlation dimension ($D$) decreases monotonically with the dataset number. The middle plot in Fig. 17 depicts such a noisy decrease in Kolmogorov entropy ($K$) vs dataset number, so that essentially no correlation exists. The first minimum of the mutual information

function (bottom plot of Fig. 17) is constant at eight time steps, rising briefing to nine for dataset #7. Thus, only the correlation dimension has a clear relation with the increasing severity of the fault in this test sequence.

A systematic search determined the set of phase-space reconstruction parameters that are most sensitive to the condition change for this test sequence. Figure 18 shows the corresponding results with a nearly linear rise in all four of the (non)connected-phase-space dissimilarities. The dissimilarity for the first dataset is not shown, because it's always zero, due to comparing the basecase with itself. The (non)connected phase-space dissimilarities clearly show the progressively worsening severity of the unbalance fault.

## 1.5 ANALYSIS OF MOTOR DATA FROM PSU

The Pennsylvania State University (PSU) operates the Applied Research Laboratory (ARL). Their facilities include the Mechanical Diagnostics Test Bed (MDTB), a motor-drive-train-generator test stand, as shown in Fig. 19 (upper picture). The gearbox is driven at a set input speed using a 30 hp, 1750 rpm AC (drive) motor, and the torque is applied by a 75 hp, 1750 rpm AC (absorption) motor. The maximum speed and torque are 3500 rpm and 225 ft-lbs respectively. The speed variation is accomplished by varying the frequency to the motor with a digital vector drive unit. The variation of the torque is accomplished by a similar vector unit capable of controlling the current output of the absorption motor. The system speed and torque set points are produced by analog input signals (0–10 VDC) supplied by the Data Acquisition (DAQ) Computer and a D/A board. The MDTB is highly efficient because the electrical power that is generated by the absorber is fed back to the driver motor. The mechanical and electrical losses are sustained by a small fraction of wall power. The MDTB has the capability of testing single and double reduction industrial gearboxes with ratios from about 1.2:1 to 6:1. The gearboxes are nominally in the 5–20 HP range. The system is sized to provide the maximum versatility to speed and torque settings. The motors provide about 2–5 times the rated torque of the selected gearboxes, and thus the system can provide good overload capability. The use of different reduction ratios and gearboxes than listed above is possible if appropriate consideration to system operation is given. The motors and gearbox are hard-mounted and aligned on a bedplate. The bedplate is mounted using isolation feet to prevent vibration transmission to the floor. The shafts are connected with both flexible and rigid couplings. Torque limiting clutches are used on both sides of the gearbox to prevent the transmission of excessive torque as could occur with gear jam or bearing seizure. In addition, torque cells are used on both sides of the gearbox to directly monitor the efficiency and the loads transmitted. Appendix B provides further test details.

The test protocol involved: (i) a linear increase in motor speed from zero to the operational speed, (ii) maintenance of this constant speed for some chosen period, (iii) a linear decrease in motor speed from (ii) to zero, (iv) sampling of the gearbox oil for particulates as a measure of gear wear. The cycle of (i)–(iv) was repeated many times under the test load until the gearbox failed, meaning that one or more gear teeth broke, causing the experiment to stop. The present analysis uses only the flattop data from (ii) of the test cycle, for which 44 datasets were available for each data channel, each containing 10 s of test data. Measurements included acceleration, acoustic sounds, temperature, torque, rotation speed, input and load power, and RMS currents. Much of this data (temperature, relative oil dielectric) are sampled much too slowly (e.g., 1 Hz) to be useful for the present analysis. Other data (rotational speed, drive power, load power, drive current, load current, and coil temperatures) were sampled at 1 kHz, but fail the data quality check also due to inadequate sampling speed. The present analysis focuses on load torque, which was sampled at 1 kHz and which passes all of the data quality tests. Thus, each 10 s dataset has 10,000 data points, all of which were concatenated serially into a single data file for ease of analysis. Figure 19 (lower plot) shows a one second sample of typical load torque data, which is quasi-periodic with complex, nonlinear features. Figure 20 shows the erratic variation over the test sequence of the linear measures of this load torque data. Skewness (solid curve in the second subplot down) decreases gradually (but very irregularly) from 0.3 to –0.2, and is the only linear measure that shows any systematic change. The first zero in the autocorrelation function (bottom subplot of Fig. 20) rises in datasets #26–27 and 30, with the failure occurring in dataset #44. Figure 21 displays the traditional nonlinear measures of the load torque, none of which has any systematic trend as the test sequence progresses. Figure 22 displays the PSDM for this test sequence, with low values (below 0.875) for datasets #1–29. All four dissimilarity measures rise abruptly, beginning at dataset #30, and remain above 0.894 for the remainder of the test sequence. These results show that only the PSDM provide forewarning of the impending failure.

## 1.6 ANALYSIS OF MOTOR DATA FROM PSU (RUN33)

ORNL decided to engage PSU (Dr. Karl Reichard) under subcontract to acquire test data for this project, based on the data analysis in the previous section. This test sequence was designated "RUN33," and also involved the MDTB, as described in App. B. Measurements included acceleration, acoustic sounds, temperature, torque, rotation speed, input and load power, and RMS currents. ORNL chose to focus on the tri-axial acceleration, based on analysis of several initial PSU sample datasets. ORNL obtained the data electronically via FTP from a passworded-account on a PSU server. Each data file contained test data for one accelerometer channel. The data were obtained at 10 min intervals through the test sequence, sampled at 102.4 kHz. The total amount of data was 4.5 GB (three accelerometer channels, times 401 snapshots for a total of 1203 files) in MatLab binary format. ORNL serially concatenated 100,000 data points from each of the data files into a single three-channel dataset for ease of analysis (1.6 GB). Each 100,000-point snapshot was divided into ten 10,000-point subsets for this analysis; the results were then averaged over these 10 cutsets to obtain a typical value for the entire snapshot. Figure 23 shows typical data for each accelerometer channel with quasi-periodic, complex, nonlinear features. Figure 24 shows conventional statistical measures throughout the RUN33 test sequence. The top plot shows noisy, gradual increases in the magnitudes of the minimum ($A_n$),

maximum ($A_x$), absolute average deviation ($a$), and sample standard deviation ($\sigma$) for each acceleration signal. The other conventional statistical measures in Fig. 24 show no clear trend through the test sequence. Figure 24 also reveals two datasets (#119 and 266) with abruptly lower signal, where the operator did not stop the data acquisition system during diagnostic shutdowns of the test. This signal loss does not invalidate the remainder of the data, so we ignore these two datasets in subsequent analysis.

The use of three orthogonal accelerations has a unique and important advantage, as follows. Acceleration is a three-dimensional vector, $\underline{A}$, that can be integrated once in time to give velocity vector, $\underline{V} = \int \underline{A}\, dt$. Mass times acceleration (vector) is force vector, $\underline{F} = m\underline{A}$. The vector dot-product of force and velocity is power (scalar), $P = \underline{F} \bullet \underline{V}$. Thus three-dimensional acceleration data can be converted into a scalar power via straight-forward methods of calculus and mechanics. ORNL previously used this approach to analyze tri-axial acceleration data from a metal cutting (lathe) operation. That analysis found that the resulting three-dimensional accelerometer power captured the relevant dynamics and had more information about the process than any single accelerometer channel. We used this same approach to obtain power using the tri-axial accelerometer data from the PSU MDTB test sequence. Figure 25 shows a sample of the RUN33 power data, which displays very complex, nonlinear features. Figure 26 shows conventional statistical measures of the RUN33 accelerometer power data. The top plot shows noisy, gradual increases in the magnitudes of the minimum ($P_n$), maximum ($P_x$), absolute average deviation ($a$), and sample standard deviation ($\sigma$) for the accelerometer power. The other conventional statistical measures of accelerometer power show no clear trend through the test sequence. Failure onset begins at dataset #394. Figure 27 shows the corresponding traditional nonlinear measures, none of which show forewarning of the failure. We conclude that only a few of conventional statistics (minimum, maximum, $a$, and $\sigma$) provide forewarning trends, but none of the traditional nonlinear measures provide indication of the impending machine failure.

In sharp contrast to the previous unclear trends, Fig. 28 shows a systematic rise in all four renormalized measures of dissimilarity, with an additional abrupt rise at the onset of failure. We obtained this result by constructing a composite measure, $C_i$, of condition change, namely the sum of the four renormalized measures of dissimilarity for each of the datasets in the test sequence. This approach shows both the rising dissimilarity as the test progresses and an indication of failure onset. We used the following algorithm to obtain this result. First, we construct the composite measure for the i-th dataset:

$$C_i = U(\chi^2) + U(\chi_c^2) + U(L) + U(L_c). \tag{1}$$

Second, we fit $C_i$ to a straight line via least-squares over a window of m datasets (#194-393 in this case):

$$y_i = ai + b. \tag{2}$$

Third, we obtain the variance of $C_i$ about the straight-line fit from step 2:

$$\sigma_l^2 = \Sigma_i (y_i - C_i)^2/(m\text{-}1). \tag{3}$$

Fourth, we determine the $\chi^2$-statistic from this straight-line fit for datasets #394-400:

$$\chi^2 = \Sigma_i (y_i - C_i)^2/\sigma_l^2. \tag{4}$$

Fifth, we maximize the value of $\chi^2$ from step 4 over the parameters ($d$, $S$, $\lambda$). The variance, $\sigma_l^2$, in step 3 measures the variability of $C_i$ about the straight-line fit over the window of m datasets (#194-393). The statistic, $\chi^2$, in step 4 measures the variability of datasets #394-400 from the straight-line fit. The value from step 4 is $\chi^2$ = 180.42, which is inconsistent with a normal distribution for 7 degrees of freedom, and is a strong indication of the failure onset. Indeed, Fig. 30 shows a clear statistical indication of failure onset. The bottom plot (labeled "normal distribution") in Fig. 29 depicts the maximum value of the $\chi^2$ statistic for n sequential values out of 200 samples from a guassian (normal) distribution with zero mean and a unity sample standard deviation. The middle curve in Fig. 29 is the maximum value of the $\chi^2$ statistic, using step 4 above, for n sequential values of the composite measure, $C_i$, over the window of $m$ = 200 datasets that span the straight-line fit (datasets #194–393). This middle curve is inconsistent with the normal distribution (bottom curve). The upper curve in Fig. 29 is the $\chi^2$ statistic, also using step 4 above, for $n$ sequential values from datasets #394–400. This upper curve (labeled "failure onset") deviates markedly from the lower curves after two datasets (#394–395), with overwhelming indication for three and more datasets. We conclude that the PSDM provide consistent indication of condition change, as well as clear indication of the failure onset.

## 1.7 ANALYSIS OF MOTOR DATA FROM PSU (RUN34)

PSU repeated the MDTB experiment (RUN34) with the same experimental parameters as RUN33. The sampling frequency was 51.2 kHz, based on our findings from RUN33. The raw data consisted of 560 datasets in MatLab binary format for each of the three orthogonal accelerations for a total of 3.2 GB, which PSU sent to ORNL on a DVD disk. ORNL serially concatenated 150,000 data points from each of the data files into a single three-channel dataset (2.1 GB) for ease of analysis. The tri-axial acceleration data was converted into one channel of accelerometer power, as described above. Each 150,000-point snapshot was divided into ten 15,000-point subsets for analysis; the results were then averaged over these 10 subsets to obtain a typical value for the entire snapshot. Data quality analysis showed that the 51.2 kHz sampling frequency was adequate. However, the experimental record of RUN34 showed numerous annotations of bad datasets, due to known experimental problems. Omission of these bad datasets reduced the total number of useful datasets by 37% to 355. In addition, erratic and inconsistent variations existed in the data amplitude. RUN34 was rejected from further analysis, due to these quality problems.

## 1.8 ANALYSIS OF MOTOR DATA FROM PSU (RUN 35)

PSU repeated the MDTB experiment for a third time (RUN35) with the same experimental parameters as RUN33. The sampling frequency was 51.2 kHz. The raw data consisted of 845 datasets in MatLab binary format for each of the three orthogonal acceleration directions, which PSU provided on a DVD disk (5GB). ORNL serially concatenated 200,000 data points from each of the data files into a single three-channel dataset for ease of analysis. The tri-axial acceleration data was converted into one channel of accelerometer power, as described above (1.8 GB). Figure 30 shows sample plots of this quasi-periodic, complex data. Each 200,000-point snapshot was divided into ten 20,000-point subsets for analysis; the results were then averaged over these 10 subsets to obtain a typical value for the entire snapshot. Data quality analysis revealed datasets with a zero signal or abrupt shifts, not unlike the previous test sequences. Consequently, we eliminated these bad datasets, as indicated by gaps in the Fig. 31 curves. We further note from Fig. 31a that only the monotonic rise (fall) in the maximum (minimum) values of power give clear forewarning of the failure onset, beginning with dataset #838. Figure 32 shows the corresponding traditional nonlinear measures. Correlation dimension and Kolmogorov entropy provide no forewarning of failure. The monotonic fall in the first minimum of the mutual information function before failure is unique and appears to be a valid (but weak) forewarning of the failure.

The top four subplots of Fig. 33 display the PSDM, using the same phase-space parameters as PSU Run33 for a demonstration of repeatability. We find a systematic rise in the PSDM, beginning at dataset #480, with an additional abrupt monotonic rise at the onset of failure, beginning at dataset #820. We obtained this result by constructing the same composite measure, $C_i$, of condition change, using the five-step process of Eqs. (1)–(4). We chose a window of 100 cutsets for Eq. (2) over datasets #720–819. We computed the $\chi^2$ statistic, beginning with dataset #820. As before, Fig. 34 depicts the maximum value of the $\chi^2$ statistic for n sequential values out of 100 samples from a guassian (normal) distribution with zero mean and a unity sample standard deviation. The middle curve in Fig. 34 is the maximum value of the $\chi^2$ statistic, using step 4 above, for n sequential values of the composite measure, $C_i$, over the window of $m$=100 datasets that span the straight-line fit over datasets #720–819. This middle curve is inconsistent with the normal distribution (bottom curve). The upper curve in Fig. 34 is the $\chi^2$ statistic, also using step 4 above, for $n$ sequential values from datasets #820–845. This upper curve (labeled "failure onset") deviates markedly from the lower curves after six datasets, with overwhelming indication for seven or more datasets. We conclude that the phase-space dissimilarity method gives clear indication of condition change, as well as forewarning of failure.

## 1.9 ANALYSIS OF TORSION DATA FROM PSU

PSU performed a second type of experiment to show failure forewarning in turbo-machinery. Specifically, this experiment simulated the effects of shifting turbine blade oscillation frequencies as a fatigue crack develops and grows. Appendix C provides details of the experiment. Figs. C.1–C.2 show eight equally spaced threaded rods that simulate the blades, which are rotated about a horizontal axis by a DC motor. Previous PSU work on this experiment showed that adjustment of lock nut locations along the threaded rods can simulate the frequency

shift, which can be sensed and tracked by measurement of the system's torsional vibration. The present experiment extended the previous PSU work by simulating a crack at the base of one of the threaded rods via a progressively deeper 0.010 in. wide slot using electric discharge machining, in lieu of moving the nuts. Test data were acquired for each depth of cut. The data were sampled at 12.8 kHz and included: motor current and voltage, two-axis translational acceleration on one bearing pillow block, and torsional vibration of the shaft. Figure 35 illustrates a representative sample of this data at three different time scales, showing complex, quasi-periodic dynamics in all channels. Our previous experience has shown that all three acceleration axes are needed to capture the machine dynamics adequately. Consequently, the two-axis acceleration data were not pursued further. Comparison of the plots in right two columns of Fig. 35 shows that the dynamical properties of the AC-coupled motor power are different from the DC-coupled motor power. Thus, the latter data was used as the best representation of the machine dynamics for all subsequent analysis.

Figure 36 shows the conventional statistical measures of the DC-coupled motor power as a function of the progressively larger slot depth. The top plot shows no systematic change in the minimum ($P_n$), standard deviation ($\sigma$), absolute average deviation ($a$), and maximum ($P_x$). The second plot from the top shows that both skewness and kurtosis are nearly constant vs slot depth. The third plot down from the top displays a series of small rises and falls in the number of time steps per cycle vs slot depth. The bottom plot illustrates a single decrease in the first zero of the autocorrelation function, followed by a constant value thereafter. Consequently, conventional statistical measures provide no forewarning.

Plots of the traditional nonlinear measures are shown in Fig. 37. The correlation dimension (top plot) decreases from 2.4 for the first dataset, to 2.2 for the third dataset, then rises gradually over the next four datasets to 2.35. The Kolmogorov entropy (middle plot) has a decrease-increase-decrease sequence that is likewise non-predictive. The first minimum in the mutual information function is constant at 6 time steps, then decreases to 5.95 at dataset 6, then rises again to 6 at the last dataset. Thus, the traditional nonlinear measures provide no failure forewarning.

Figure 38 shows that all four PSDM rise monotonically from near zero to large values. This result was obtained by combining the four PSDM into a single composite measure, $C_i$, as described by Eq. (1). An exhaustive search then was performed over the parameter space of the number of phase-space dimensions ($d$), the number of phase-space symbols ($S$), and the time-delay lag ($\lambda$). The search revealed a single set of parameters that give the monotonic rise in $C_i$. The meaning of these parameters is explained in App. A. We conclude that the PSDM show a systematic predictive trend for the progressive failure in this experiment.

## 1.10 ANALYSIS OF BEARING DATA FROM PSU

PSU performed a third type of experiment to show forewarning of bearing failure. The Bearing Prognostic Test Rig uses a pair of double row spherical roller support bearings to support a shaft on which the test bearing is held in place by a bearing holder between the two support bearings. The test rig was configured to test deep groove ball 1 1/8 in. test bearings, but is configurable to other types of bearings. A load jack and load cell were mounted directly behind the test bearing

and have a radial load capacity of 1,000 lbs in phase with the outer ring. The system is driven by a SCR motor with an integral tachometer fitted to the motor. The motor is coupled to the test shaft with a flexible coupler to reduce transmitted vibration from the motor. Piezoelectric accelerometers are currently utilized with frequency ranges between 10 and 100KHz. Data were collected on a 48 channel National Instrument 4472 PXI system. The test ran continuously in the overloaded state, with data acquired at periodic intervals until the bearing finally failed. Appendix D provides details of the test plan and experimental setup.

We describe the data analysis next. Five channels of data were sampled at a frequency of 51.2 kHz, including tri-axial acceleration at the test bearing, as well as motor current and motor voltage. PSU provided this data to ORNL on DVD disks, in the form of MatLab™ binary files (5.6GB). ORNL serially concatenated 500,000 data points from each of the data files into a single three-channel dataset with 895 segments of three-channel accelerometer data (16.4GB). The tri-axial acceleration data were converted into one channel of accelerometer power, as described above (5.7GB); the 500,000 data points in each segment were subdivided into ten cutsets of 50,000 points each for the subsequent analysis. Figure 39 shows time-serial samples of the accelerometer power, which has a very complex, nonlinear waveform with rises and falls in envelop amplitude over 200 ms.

Figure 40 displays the conventional statistical measures of the accelerometer power. The minimum and maximum (top plot) are variable about a series of plateaus, rising and falling without a clear trend. Skewness (bottom curve in the second plot from the top) is roughly constant during the entire sequence. Kurtosis (top curve in the second plot from the top) is moderately variable during the first one hundred cutsets, more variable and slightly larger over cutsets #100–300, then smaller and much less variable for cutsets #300–810, and finally larger and more variable for cutsets #811–895. The average number of time steps per cycle (third plot from the top) decreases gradually but erratically from 16 to 13. The first zero in the autocorrelation function (bottom plot) decreases erratically from five to four over cutsets #1–300, and then is constant at four thereafter. All of these conventional statistics display an abrupt spike at cutsets #104–105. These conventional statistical measures provide no clear forewarning of failure.

Figure 41 depicts the traditional nonlinear measures of accelerometer power. Correlation dimension (top plot) is moderately variable over the entire sequence with a clear spike at cutsets #104–105. Kolmogorov entropy (middle plot) is very variable over the whole sequence without a clear trend. $M_1$ (bottom plot) falls abruptly from 4.7 to 4 at cutset #104, then remains constant at four thereafter. These traditional nonlinear measures give no failure forewarning.

The top four subplots of Fig. 42 display the four PSDM, with the same spikes as the other measures in datasets #104–105 and #200. Ignoring these spikes, the PSDM rise systematically, beginning at dataset #600. An additional abrupt rise occurs at failure onset, beginning at dataset #795. We used the five-step process of Eqs. (1)–(4) to obtain this result. We chose a window of one hundred cutsets for Eq. (2) over datasets #695–794. We computed the $\chi^2$ statistic, beginning with dataset #795. As before, the bottom curve in Fig. 43 depicts the maximum value of the $\chi^2$ statistic for n sequential values out of 100 samples from a Gaussian (normal) distribution with

zero mean and a unity sample standard deviation. The middle curve in Fig. 43 is the maximum value of the $\chi^2$ statistic, using step 4 above, for $n$ sequential values of the composite measure, $C_i$, over the window of $m$ = 100 datasets that span the straight-line fit over datasets #695–794. This middle curve is inconsistent with the normal distribution (bottom curve). The upper curve in Fig. 43 is the $\chi^2$ statistic, also using step 4 above, for $n$ sequential values from datasets #795–819. This upper curve (labeled "failure onset") provides clear indication of failure onset after three or more consecutive datasets. We conclude that the PSDM gives clear indication of condition change, as well as forewarning of failure.

## 2. DISCUSSION

The objective of this phase is a compelling demonstration of the phase-space dissimilarity approach for forewarning of failures. Thus, we studied a variety of failures and different data. Table 4 below summarizes the present results. The right-most column shows the measures that we tested for failure forewarning: conventional statistical measures [CSM] (data maximum [MX], data minimum [MN], absolute average deviation [$a$], standard deviation [$\sigma$], skewness [SK], kurtosis [KT], average number of time steps per cycle [TS]), traditional nonlinear measures [TNM] (correlation dimension [CD], Kolmogorov entropy [KE], first minimum in the mutual information function [$M_1$], and PSDM.

Table 4. Data and Failures

| Data Provider | Equipment and Type of Failure | Diagnostic Data | Measures to Forewarn of Failure |
|---|---|---|---|
| EPRI | Electric motor airgap offset | Motor power | CSM TNM PSDM |
| EPRI | Electric motor broken rotor | Motor power | PSDM |
| EPRI | Electric motor turn-to-turn short | Motor power | PSDM |
| Otero (Spain) | Electric motor imbalance | 1D acceleration | MN α σ MX KT CD PSDM |
| PSU/ARL | Overloaded gearbox | Load torque | SK PSDM |
| PSU/ARL | Overloaded gearbox (RUN 33) | Accelerometer power | MN α σ MX PSDM |
| PSU/ARL | Overloaded gearbox (RUN 35) | Accelerometer power | MN MX $M_1$ PSDM |
| PSU/ARL | Crack in rotating blade | Motor power | PSDM |
| PSU/ARL | Motor-driven bearing | Accelerometer power | PSDM |

Only the PSDM provide forewarning of the machine failures across all of these test sequences. This forewarning indication is present for several different kinds of equipment failures, as well as for different types of diagnostic data. Moreover, this work developed a statistical criterion for the determination of failure onset, based on the sum-of-squares deviation from a straight-line fit to the rising trend in PSDM. A goal for the third project year is extension of this statistical criterion into a statistical test for forewarning of failure. A second goal is publication of this work in a peer-reviewed technical journal.

### 2.1 LESSONS LEARNED

Data exchange during this second project year typically involved >10MB for each test sequence. Consequently, we were unable to use e-mail attachments, as in the first project year. We found FTP (file transfer protocol) transfers were slow (many hours), and frequently failed due to network or server problems. The most reliable method consisted of compressing the data into ZIP-files, writing the files to CD-ROM or DVD disks, and sending the disks via over-night mail. This latter method was used for most of the PSU/ARL data, and all of the EPRI data.

A low data sampling rate was used for acquiring some of the data (Otero/Spain and the first PSU test sequence). Consequently, the PSDM displayed undesirable variability, but still provided forewarning of the failures. ORNL obtained several trial datasets from PSU after initiating their subcontract to assure that the data quality (especially the data sampling rate) was adequate for our analysis. ORNL analysis of these sample data showed that a sampling rate of >50 kHz provided adequate data quality. Moreover, this analysis determined that current and voltage from a digitally controlled electric motor was not appropriate for PSDM analysis. For such tests, we focused instead on 3D accelerometer power.

One lesson-learned from the first project year of this work is that a higher data-sampling rate (>50 kHz) is needed for the nonlinear analysis. Tests during the first project year used a PdMA Emax system, which is a ruggedized laptop computer for data acquisition with a maximum sampling rate of 12,288 Hz for each of six channels. Consequently, ORNL procured hardware and software from National Instruments Corporation during the first quarter of this second project year with the capability of sampling each of six channels at 208 kHz. The system components include a data acquisition board (PCI-MIO-16E-1), connector block, cable, and companion software (LabView™) at a total cost of $4,180. Table 5 provides further details. ORNL installed and tested this data acquisition system on the desktop PC that was procured for this project during PY1.

## 2.2 OTHER WORK

Dennis Strickler of Computational Science and Engineering Division worked with Lee Hively to port the nonlinear analysis code to ORNL's 184-node IBM SP computer, called "Eagle." Code changes used the Message-Passing Interface (MPI) for parallelization. Measures of dissimilarity for each channel of sample data were computed independently on multiple processors. Initial results gave an improvement in computational time by a factor of three. Further improvements are possible by parallelizing the base case calculations and implementing shared memory parallel programming techniques in individual modules.

We submitted a patent application (PA) to the U.S. Patent Office (USPO) on March 8, 2000, seeking protection of the connected phase-space dissimilarity approach. The USPO issued an office action on January 16, 2002, which was received by the ORNL legal department on February 4, 2002. We sent our response to the USPO on May 15, 2002. The attorney of record on this response was a subcontractor, Michael McGovern of Quarles and Brady LLP (Milwaukee, Wisconsin). The USPO responded in July 2002 with a notice that allowed all of our revised claims.

Our research has developed improvements to the phase-space dissimilarity methodology, as described in a more recent invention disclosure (ERID #0885), dated October 17, 2000. This ID was elected for conversion into a PA on December 19, 2000. We have added further improvements to the methodology, and submitted the PA to the USPO on July 12, 2002, also in collaboration with Michael McGovern.

**Table 5. Further Details of New Data Acquisition System**

| Brief Description of Item | Cost ($) | Additional Details |
|---|---|---|
| 1) PCI-MIO-16E-1 Multichannel IO Board (part number 777305-01) | 1,795 | Resolution = 12 bits<br>Sampling Rate = 1.25 MHz total or 208 kHz for each of 6 channels<br>Voltage range = +/- 10V maximum or +/- 0.05V minimum<br>Absolute accuracy < 0.045%<br>Input impedance = 100 Gohms in parallel with 100pF<br>Warmup time = 15 min<br>Calibration interval = 1 year<br>Operating temperature = 0 – 55°C<br>Storage temperature = -20–70 °C<br>Relative humidity = 10–90% noncondensing |
| 2) SH68-68-EP (1 meter shielded cable) (part number 184749-01) | 95 | Provide connection from item #1 to item #3 |
| 3) SCB-68 Shielded IO connector block (part number 776844-01) | 295 | Provide connection points for current and voltage probes |
| 4) LabView™ full development software (part number 776670-03) acquisition/analysis | 1,995 | For Windows 2000/NT/Me/9x (English)<br>Graphical user interface and programming environment for data |
| **TOTAL COST** | 4,180 | |

ORNL received an e-mail request on June 11, 2002 from Marty Martinez (Enegetics, Inc.) for a brief annual report to be included in an FY02 NERI program summary. We prepared two pages of text plus one figure in the prescribed format and sent it to Martinez via e-mail on June 18, 2002. We received a draft PDF version of our summary from Martinez via e-mail on August 7, 2002 for our comments, which we provided to him on August 12, 2002 also via e-mail.

ORNL's nonlinear paradigm is applicable to forewarning of biomedical events, as well as forewarning of machine failures. Specifically, we are using the phase-space dissimilarity approach to analyze scalp EEG to forewarn of epileptic seizures. Recent results are documented in a paper to the journal, *IEEE Transactions on Biomedical Engineering*, for a focus issue on epileptic seizure prediction. We received the reviewers' comments on July 15, 2002. We revised the paper in accord with the reviewers' suggestions, and sent the revised paper and our response to TBME on August 6, 2002. Since this work is presently unfunded, we recently submitted proposals to two potential sponsors. We have also submitted two proposals to analyze EKG data for forewarning of cardiac events.

Lee Hively was contacted by Diana Tallett (Program Administrator for the Nuclear Safety and Technology Product Line at PNNL) on October 8, 2001, regarding the total dollar value of the present NERI project. Since the abstract for this work is available on the NERI Web site (neri.ne.doe.gov), Hively provided the total amount ($1.117M) to Ms. Tallett, who confirmed receipt of this information in an October 9, 2002 e-mail.

On March 28, 2002, Lee Hively was contacted by Richard Wood (in ORNL's Nuclear Science and Technology Division), who is a collaborator on a different NERI project. Wood requested a short (two paragraph) summary of our work for the Nuclear Regulatory Commissioners, which we provided that same day.

## 2.3 ISSUES/CONCERNS

Lee Hively's PII-400MHz PC failed late in the afternoon of Friday, July 12, 2002, and was diagnosed on Monday, July 15, 2002 with a harddrive crash. This computer was approaching four years old, and consequently was replaced with a new computer under ORNL's Managed Hardware Program (procurement initiated on July 16, 2002 with installation on July 30, 2002). The total cost of this procurement was $4,813, plus $1,300 for an upgrade to MatLab™ (version 6.5) and $384 for an upgrade to Compaq Visual FORTRAN™ (version 6.6). Hively's second office PC was inadequate for compute-intensive analysis and had been out of commission since May 2002 for upgrades. The first motherboard/CPU upgrade failed repeatedly and reproducibly when loaded with a compute-intensive problem. Replacement hardware failed in the same way, and also was returned to the vendor. The second set of replacement hardware was provided by PC support to Hively on July 18, 2002, and successfully reproduced sample results. No subsequent failures have occurred. The total cost of this upgrade was $6,108. These failures did

not impact the work progress, because in the intervening two weeks (July 15–30, 2002), Hively used a PC at home (2GHz P4-Xeon™, 1GB memory, two 73GB SCSI harddrives) for data

analysis via Compaq™ Visual FORTRAN, visualization via MatLab™, and report preparation via MS-Word™.

## 2.4 COST PERFORMANCE

We received $157,000 for the first project year and $481,000 for the second project year. Total project spending through the fourth quarter of the second project year (through the fiscal month of August 2002) is $580,761. Our work is on schedule and within the budget, as shown in Fig. 44 and Table 6.

Subcontract work by Dr. Karl Reichard (Applied Research Laboratory, Pennsylvania State University) has proceeded within budget ($90K) and on schedule by providing test sequence data to ORNL. The FY02 subcontract payments to PSU were as follows: $8,376.24 on March 19, 2002, $22,336.46 on April 25, 2002, $15,544.31 on May 16, 2002, $19,742.81 on June 19, 2002, and $23,974.18 on July 15, 2002 for a total of $89,974.00. ORNL is pleased with PSU's performance, and plans to use them for FY03 work for this project.

During the first project year, we demonstrated the PSDM approach for two test sequences, costing $157K, or $78.5K/test. During this second project year, we gave a compelling demonstration of the method with nine test sequences for $481K, or $53.4K/test. These values correspond to an efficiency improvement of (78.5/53.4) – 1, or nearly 50%.

Several factors have contributed to this efficiency improvement. One innovation is the statistical criterion for failure onset (Eqs. 1–4), as an easily computable objective function for the best choice of phase-space parameters. We included this criterion in our most recent patent application. The present subcontractor (ARL/PSU) is the second contributor to this success by providing high-quality data for the test sequences that we have presented here. A third factor is the consistent use of a key lesson-learned from the first year of this project, namely requiring a sufficiently high sampling rate as a crucial parameter in the data quality. A fourth factor involves refinement of the research-class FORTRAN software implementation of the methodology by adding new high-level routines for partial automation of the PSDM analysis, rewriting modules for clear algorithmic flow, and combining related modules. A final reason for this accomplishment involves development of MatLab™ m-files to automate the search through PSDM results for many different parameter choices, and also to provide publication-quality plots of the results. These improvements led to a much more efficient use of the analyst's time.

The specific goals for the third year's work are as follows. First, we will work with our subcontractor to acquire and analyze additional test-sequence data for further demonstration of the technology for forewarning of machine failure. Second, we will work with the subcontractor to assess the impact of this technology, in terms of enhanced safety at next-generation nuclear power plants and corresponding cost reductions. Third, we will collaborate with operators at ORNL's High-Flux Isotope Reactor (HFIR) to acquire and analyze operational data for representative nuclear-grade equipment. HFIR currently experiences several failures per year in this equipment, so forewarnings of failure also would provide an immediate safety benefit to

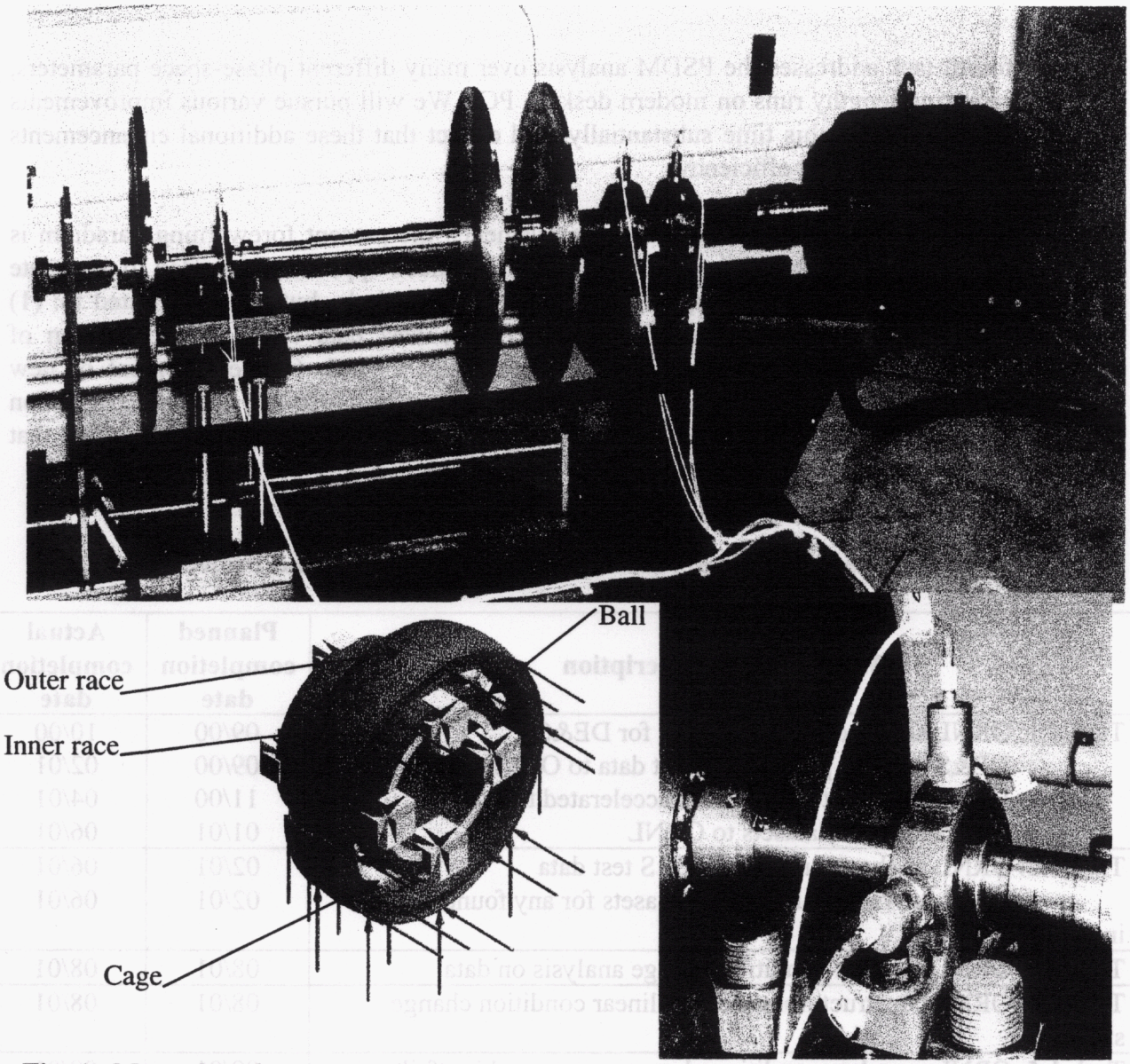



**Fig. 1. Motor-bearing system (top), bearing details (bottom left), and bear assembly (bottom right).**

HFIR. A fourth task addresses the PSDM analysis over many different phase-space parameters, presently requiring lengthy runs on modern desktop PCs. We will pursue various improvements in the analysis to reduce this time substantially, and expect that these additional enhancements will provide even better cost efficiency.

Notwithstanding past and future efficiency improvements, the present forewarning paradigm is limited by the central role of the human analyst. An advanced approach is needed to automate many different aspects of human expertise. These features include, but are not limited to: (1) assessment of data quality indicators; (2) choice of the best data type; (3) comparison of conventional statistics, traditional nonlinear measures, and PSDM; (4) development of new nonlinear measures; and (5) statistical test(s) for failure forewarning. While the implementation of these steps in fully automated form is beyond the scope of the present project, we think that such an approach should be pursued in the future.

**Table 6. Status Summary of NERI Tasks for First and Second Project Years**

| Milestone/task description | Planned completion date | Actual completion date |
|---|---|---|
| Task 1.1: ORNL set subcontract in place for DE&S | 09/00 | 10/00 |
| DE&S provide preliminary test data to ORNL | 09/00 | 02/01 |
| DE&S construct test plan for accelerated testing | 11/00 | 04/01 |
| DE&S provide datasets to ORNL | 01/01 | 06/01 |
| Task 1.2: ORNL analyze quality of DE&S test data | 02/01 | 06/01 |
| DE&S provide replacement datasets for any found inadequate | 02/01 | 06/01 |
| Task 1.3: ORNL perform condition change analysis on data | 08/01 | 08/01 |
| Task 1.4: ORNL construct library of nonlinear condition change signatures | 08/01 | 08/01 |
| Task 1.5: ORNL correlate condition change to approaching failure | 08/01 | 08/01 |
| Task 1.6: ORNL procure new computer | 08/01 | 05/01 |
| ORNL implement nonlinear analysis software on new PC | 08/01 | 06/01 |
| Task 2.1: PSU provide test data for several seeded-fault sequences | 06/02 | 06/02 |
| Task 2.2: ORNL evaluate prognostication capability of nonlinear paradigm | 08/02 | 08/02 |
| Task 2.3: ORNL improve nonlinear paradigm as appropriate | 08/02 | 08/02 |
| Task 2.4: ORNL develop algorithm for pattern change recognition | 08/02 | 08/02 |

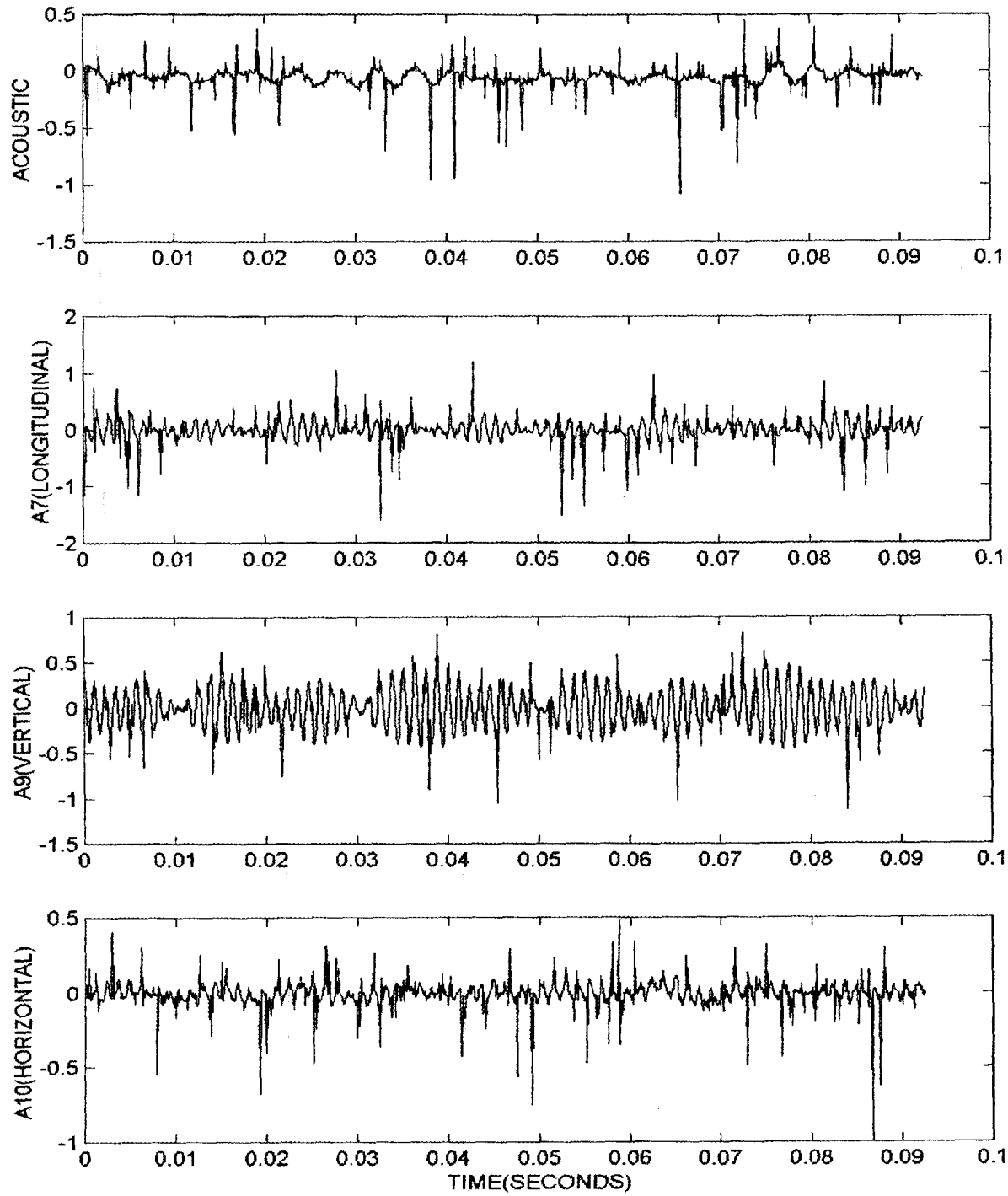



**Fig. 2. Baseline data vs time from the equipment in Fig. 1.** Top plot shows acoustic data. Second plot down shows accelerometer data. Third plot down shows accelerometer data. Bottom plot shows another channel of accelerometer data. Vertical axes in all four plots are arbitrary units.

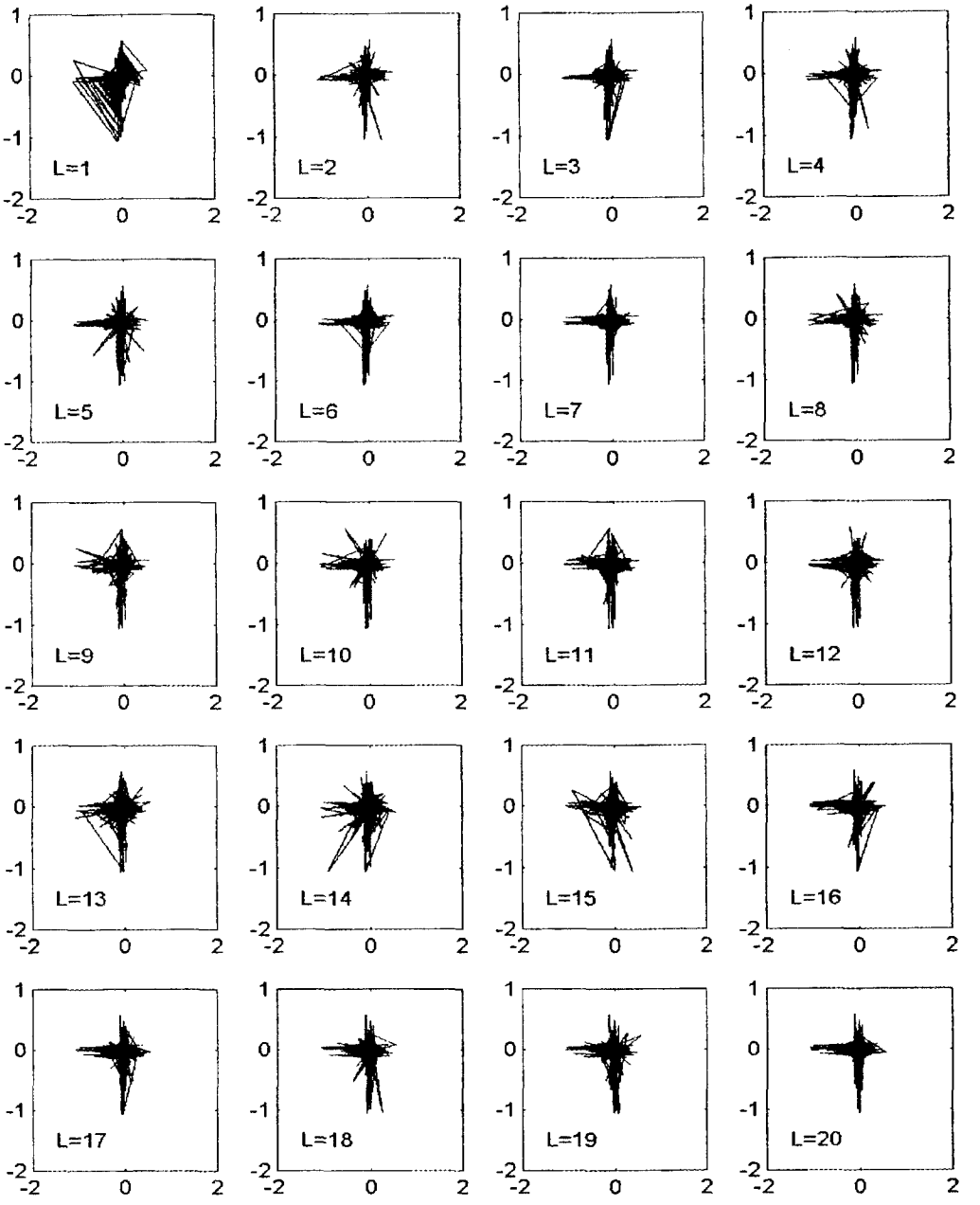



**Fig. 3. Two dimensional phase-space plots for acoustic data in top plot of Fig. 2.** The abrupt spikes in that time serial data cause the bow-tie and cross-shaped phase-space portraits. The lack of change for L≥2 indicates that this data is under-sampled.

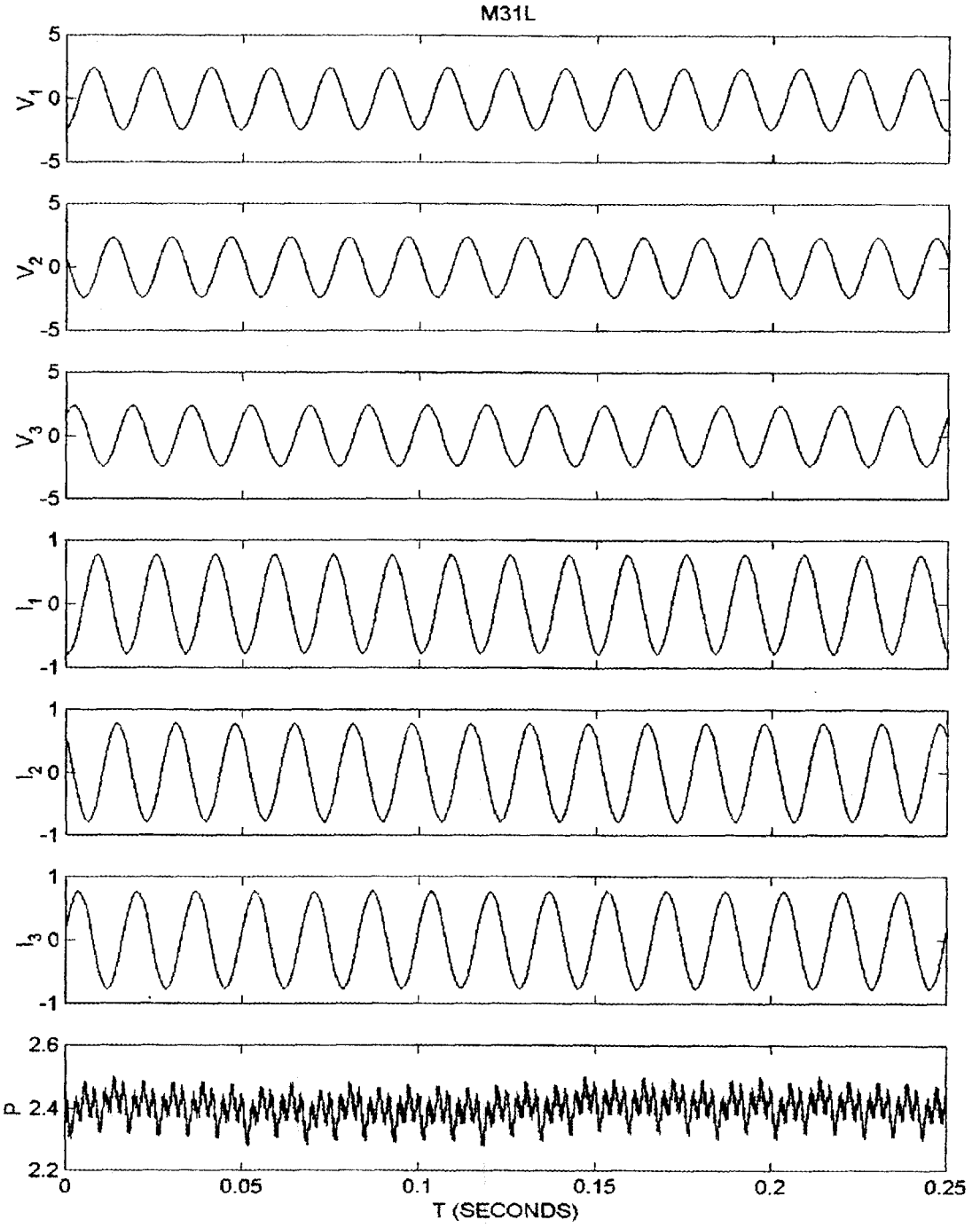



**Fig. 4. Typical baseline data vs time from the Allis Chalmers motor**. Top three plots show the three-phase voltages ($V_i$). Middle three plots show the three-phase currents ($I_i$). The sinusoidal variation in these plots corresponds to 60 Hz. The bottom plot shows instantaneous power, $P$, as the sum of the products of the three-phase currents times the corresponding voltages. See text for discussion.

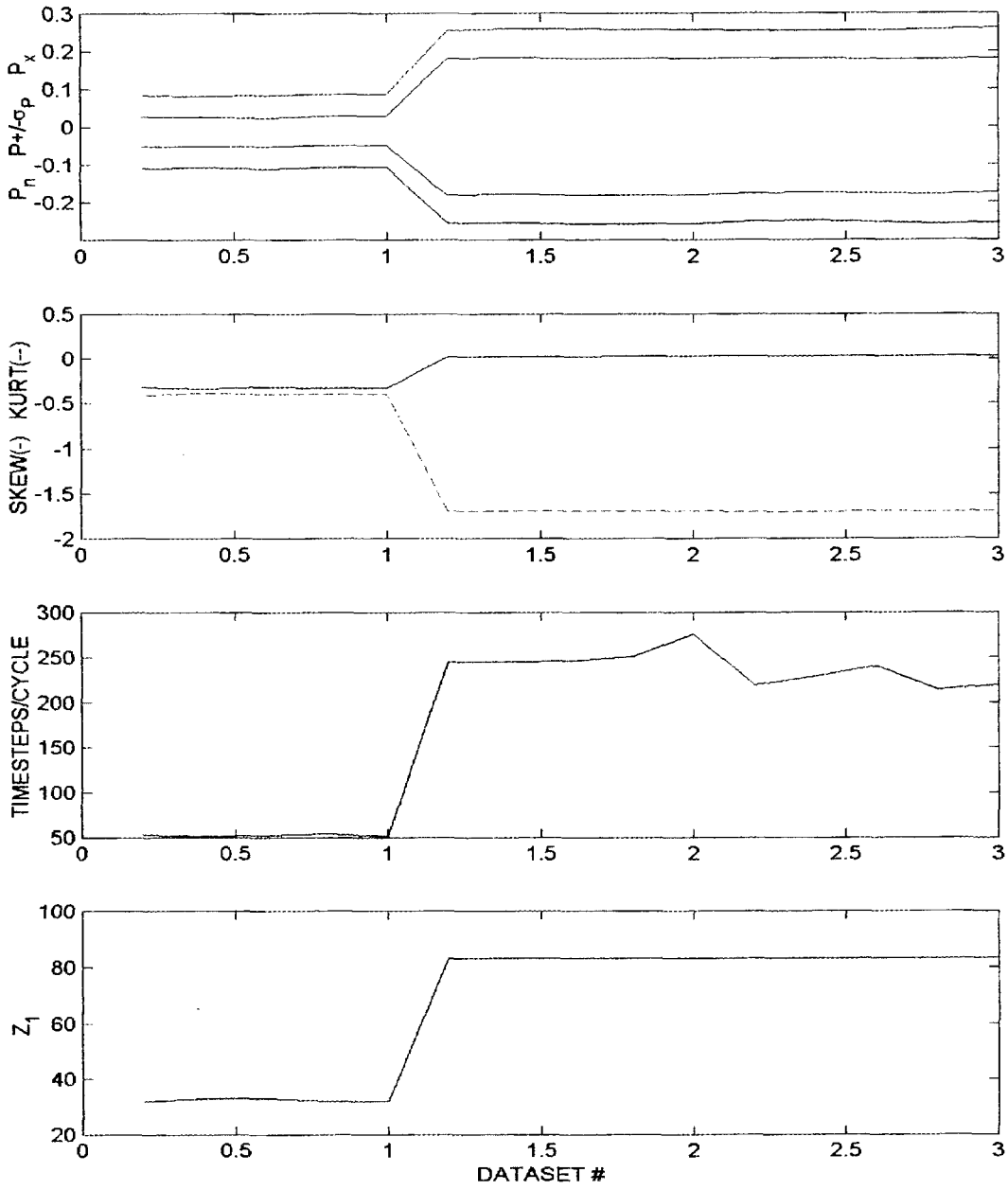



**Fig. 5. Linear measures for the air-gap seeded-fault.** Top plot shows various linear measures of instantaneous power, $P$: minimum ($P_n$) as the bottom curve, maximum ($P_x$) as the top curve, average plus one standard deviation ($P + \sigma_P$) as the middle top curve, and average minus one standard deviation ($P - \sigma_P$) as the middle bottom curve. Second plot down shows skewness (solid) and kurtosis (- -) in the instantaneous power. Third plot down shows the number of time steps per cycle in the instantaneous power. Bottom plot shows the lag in time steps, corresponding to the first zero in the autocorrelation function. See text for discussion.

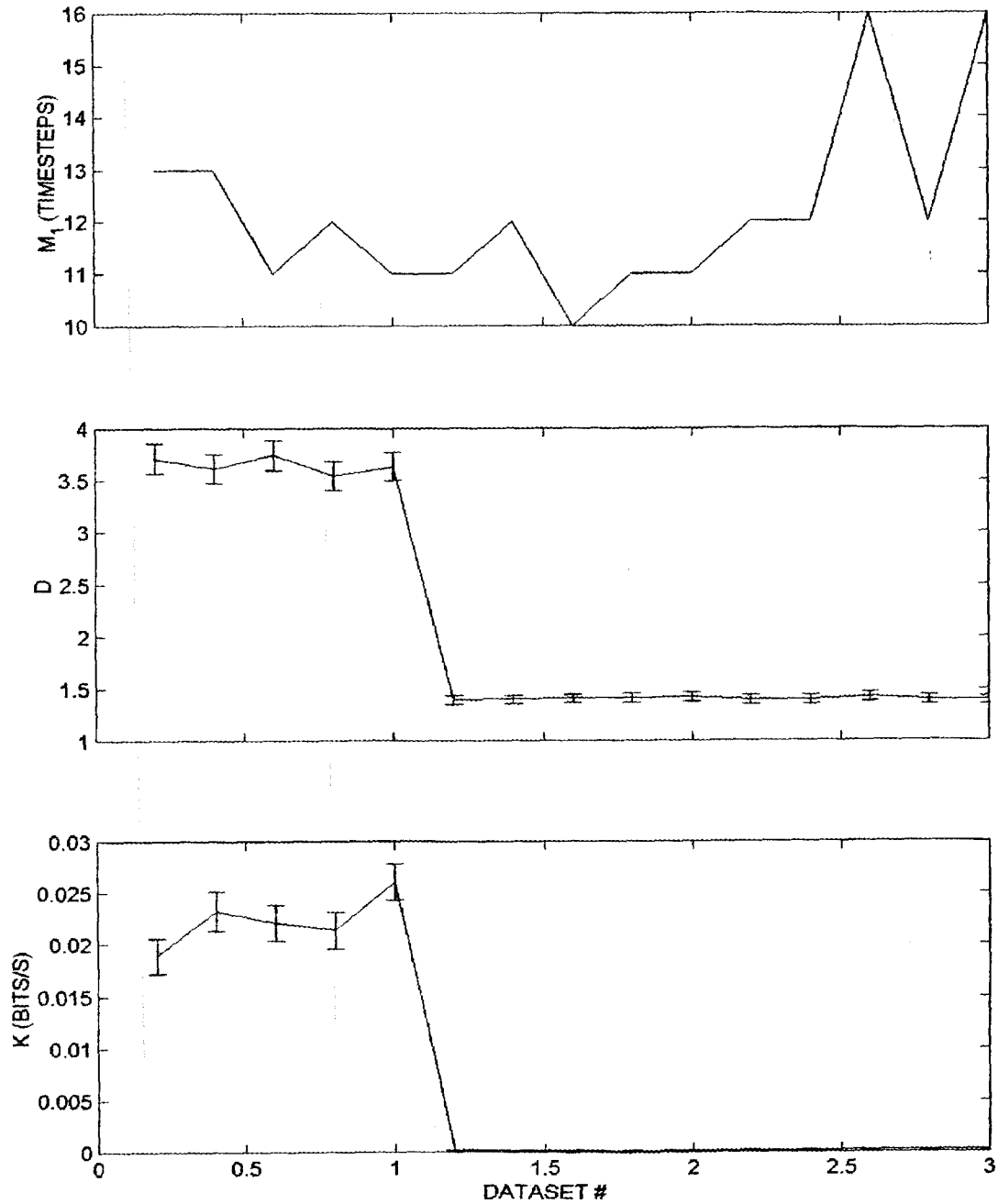



**Fig. 6. Conventional nonlinear measures for the air-gap seeded-fault.** Top plot shows the location (in time steps) of the first minimum in the mutual information function. The middle plot shows the correlation dimension (*D*). The bottom plot shows the Kolmogorov entropy (*K*). Error bars in the middle and bottom plots correspond to the 95% confidence interval. See text for discussion.

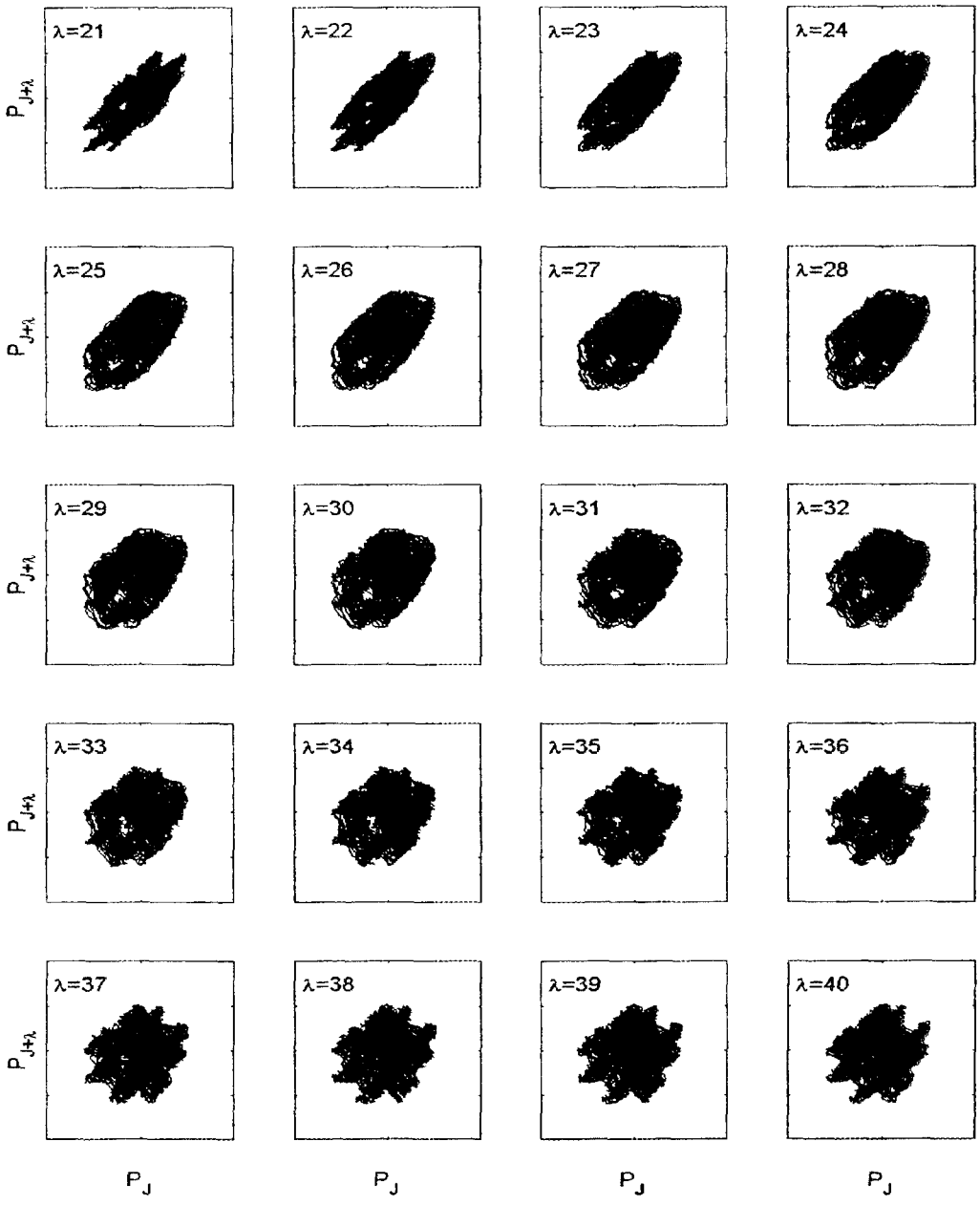



**Fig. 7. Two-dimensional phase-space reconstructions for baseline power data.** The horizontal axis is power, $P_J$, at time $t_J$. The vertical axis is power, $P_{J+\lambda}$ at later (lagged) time $t_{J+\lambda}$. Each plot corresponds to progressively larger time lag, $\lambda$, beginning at twenty-one (21) time steps in the upper left plot. The lag increases by one time step as the plots progress from left to right. The lag also increases as the plots descend from top to bottom, as shown by the "$\lambda$=nn" designation in the upper left corner of each plot, where "nn" denotes the particular value of lag. Successive plots demonstrate the "unfolding" of the phase-space representation. See text for further discussion.

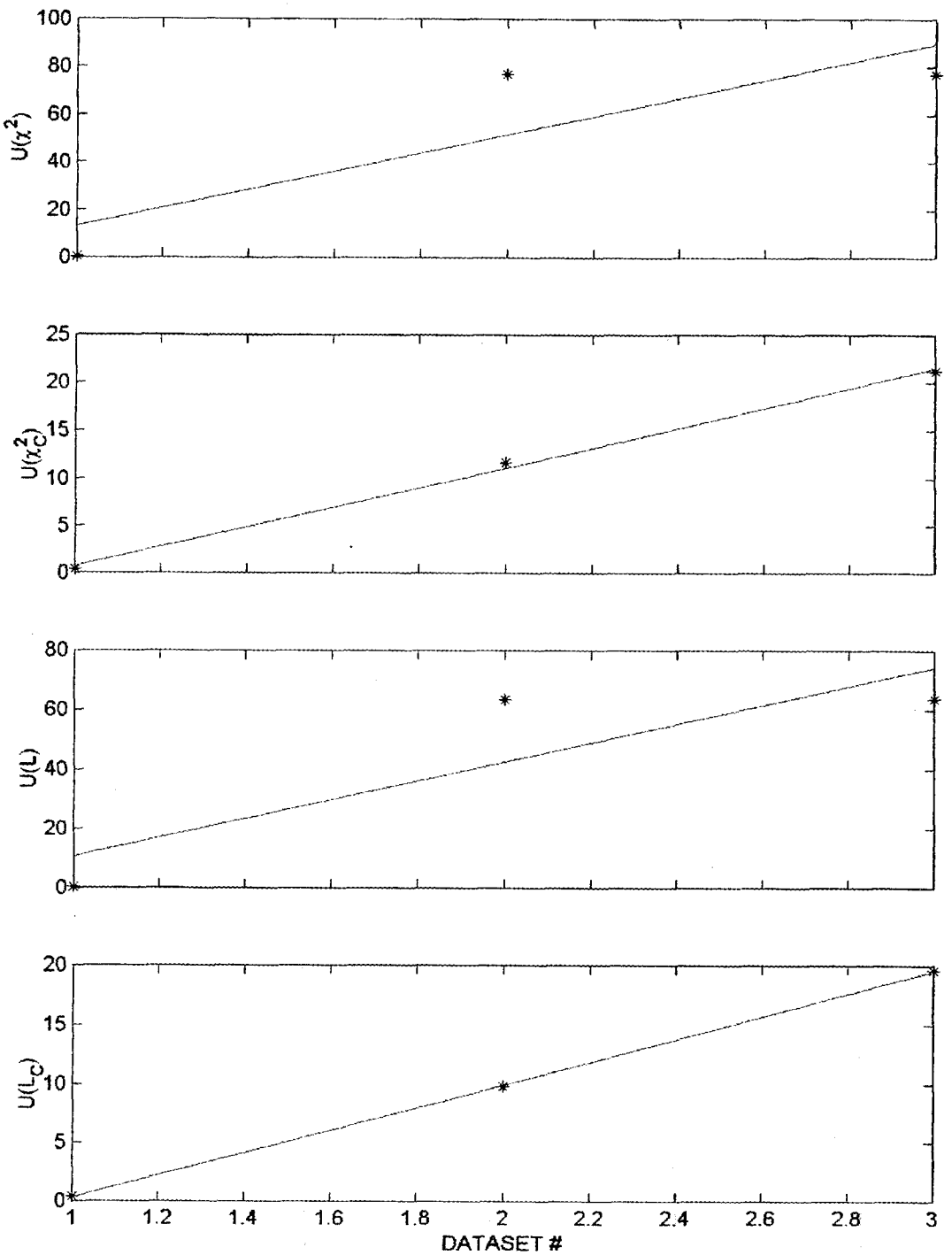



**Fig. 8. PSDM for the airgap-offset seeded-fault.** The phase-space parameters: $d$=3, $S$=56, $w$=573, $\lambda$=88. Dataset #1 is for the nominal (no fault) state. Datasets #2-3 are for two different airgap-offset faults. See text for further discussion. Stars (*) show the dissimilarity values with the straight linear added as an aid for interpretation of the graphs.

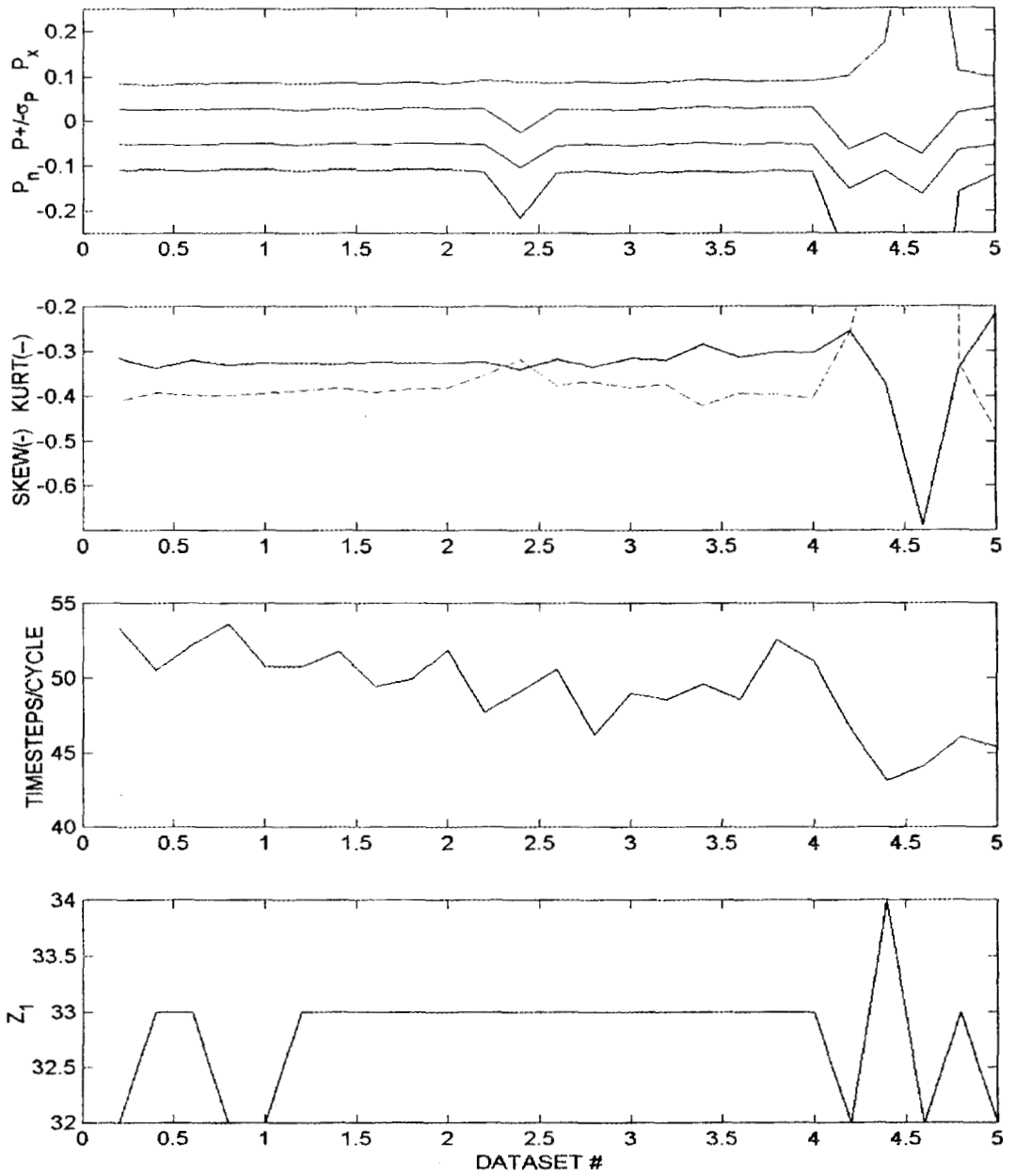



**Fig. 9. Linear measures for the broken-rotor seeded-fault.** Top plot shows various linear measures of instantaneous power, $P$: minimum ($P_n$) as the bottom curve, maximum ($P_x$) as the top curve, average plus one standard deviation ($P + \sigma_P$) as the middle top curve, and average minus one standard deviation ($P - \sigma_P$) as the middle bottom curve. Second plot down shows skewness (solid) and kurtosis (- -) in the instantaneous power. Third plot down shows the number of time steps per cycle in the instantaneous power. Bottom plot shows the lag in time steps, corresponding to the first zero in the autocorrelation function. See text for discussion.

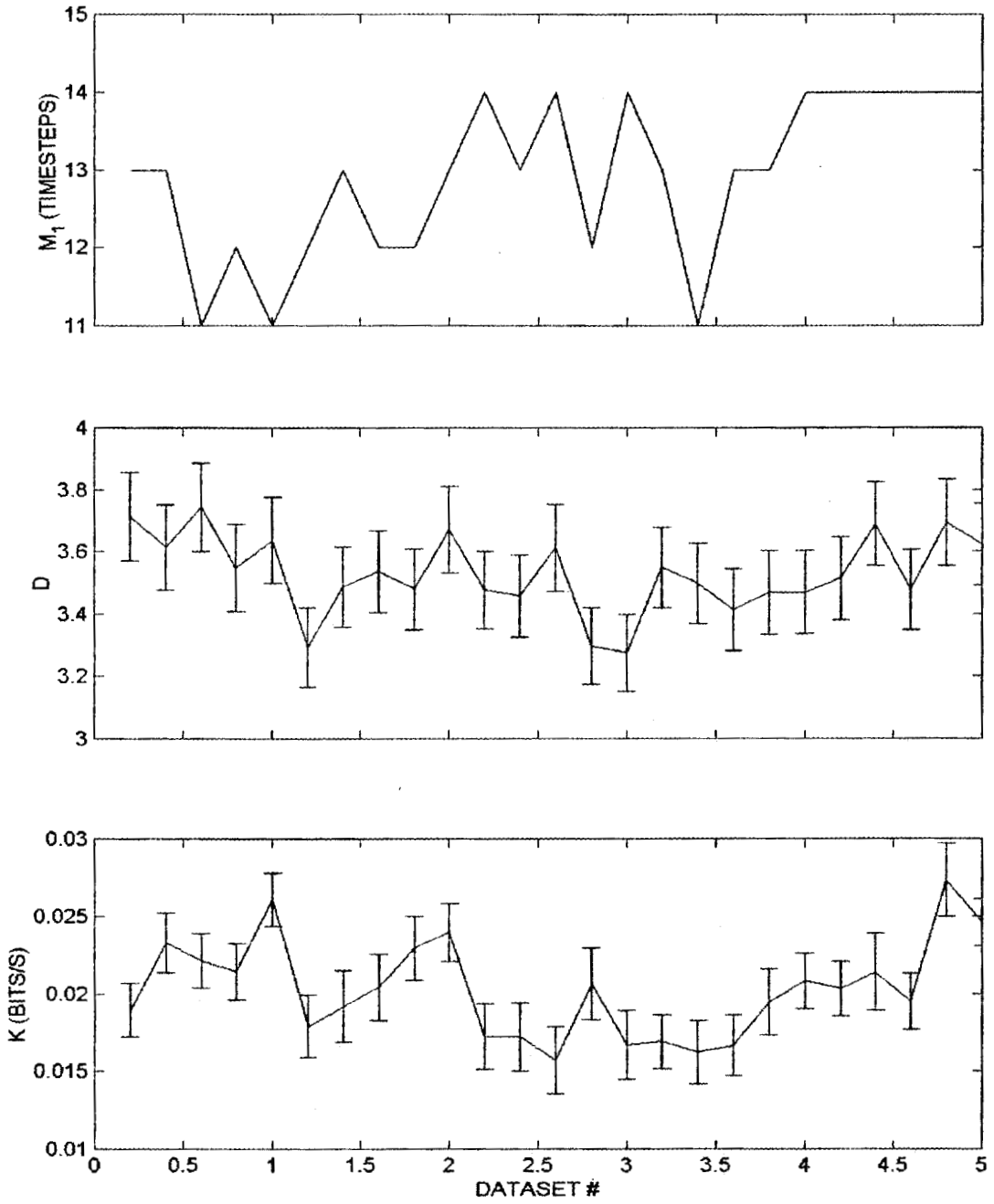



**Fig. 10. Conventional nonlinear measures for the broken-rotor seeded-fault.** Top plot shows the location (in time steps) of the first minimum in the mutual information function. The middle plot shows the correlation dimension (*D*). The bottom plot shows the Kolmogorov entropy (*K*). Error bars in the middle and bottom plots correspond to the 95% confidence interval. See text for discussion.

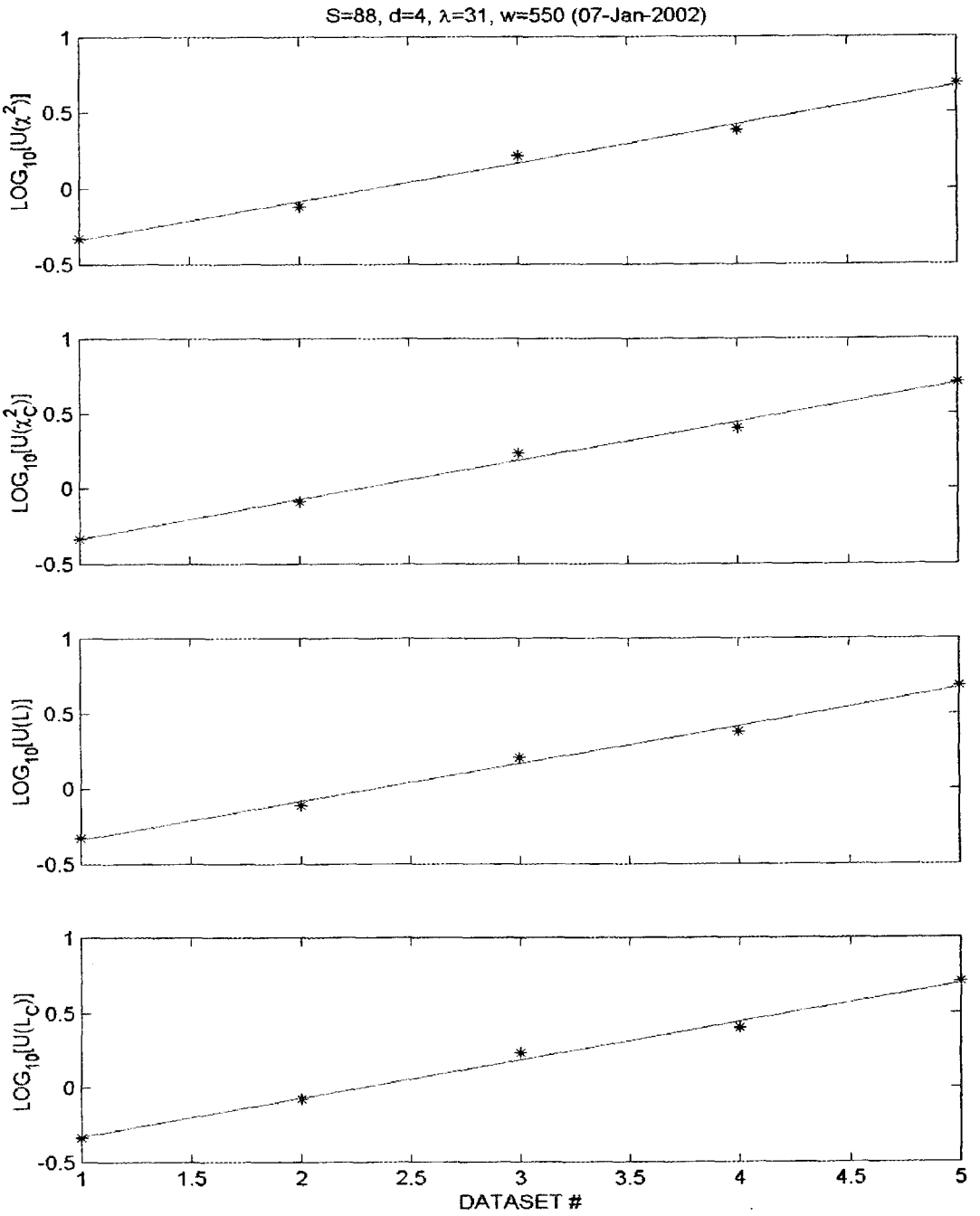



**Fig. 11. PSDM for the broken-rotor seeded-fault.** Dataset #1 is for the nominal (no fault) state. Dataset #2 is for the 50% cut in one rotor bar. Dataset #3 is for the 100% cut in one rotor bar. Dataset #4 is for two cut rotor bars. Dataset #5 is for four cut rotor bars. The exponential rise in the severity of the seeded faults is shown as an almost linear rise (solid line) in the logarithm of all four dissimilarity measures (*) for the chosen set of phase-space parameters.

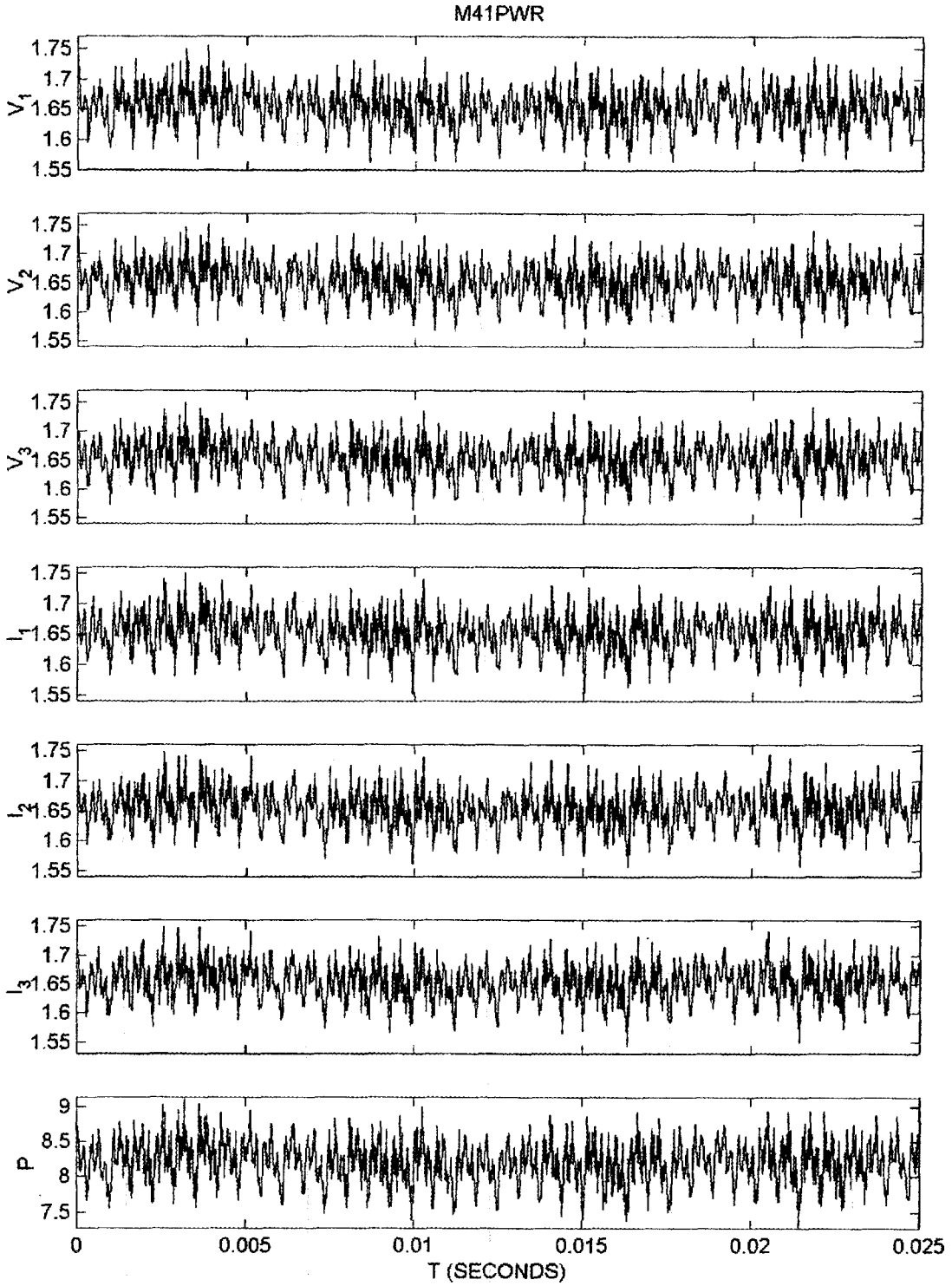



**Fig. 12. Typical baseline data vs time from the GE motor.** Top three plots show the three-phase voltages ($V_i$). Middle three plots show the three-phase currents ($I_i$). Note the very rich dynamical features in this raw data, as well as in the instantaneous power (bottom plot).

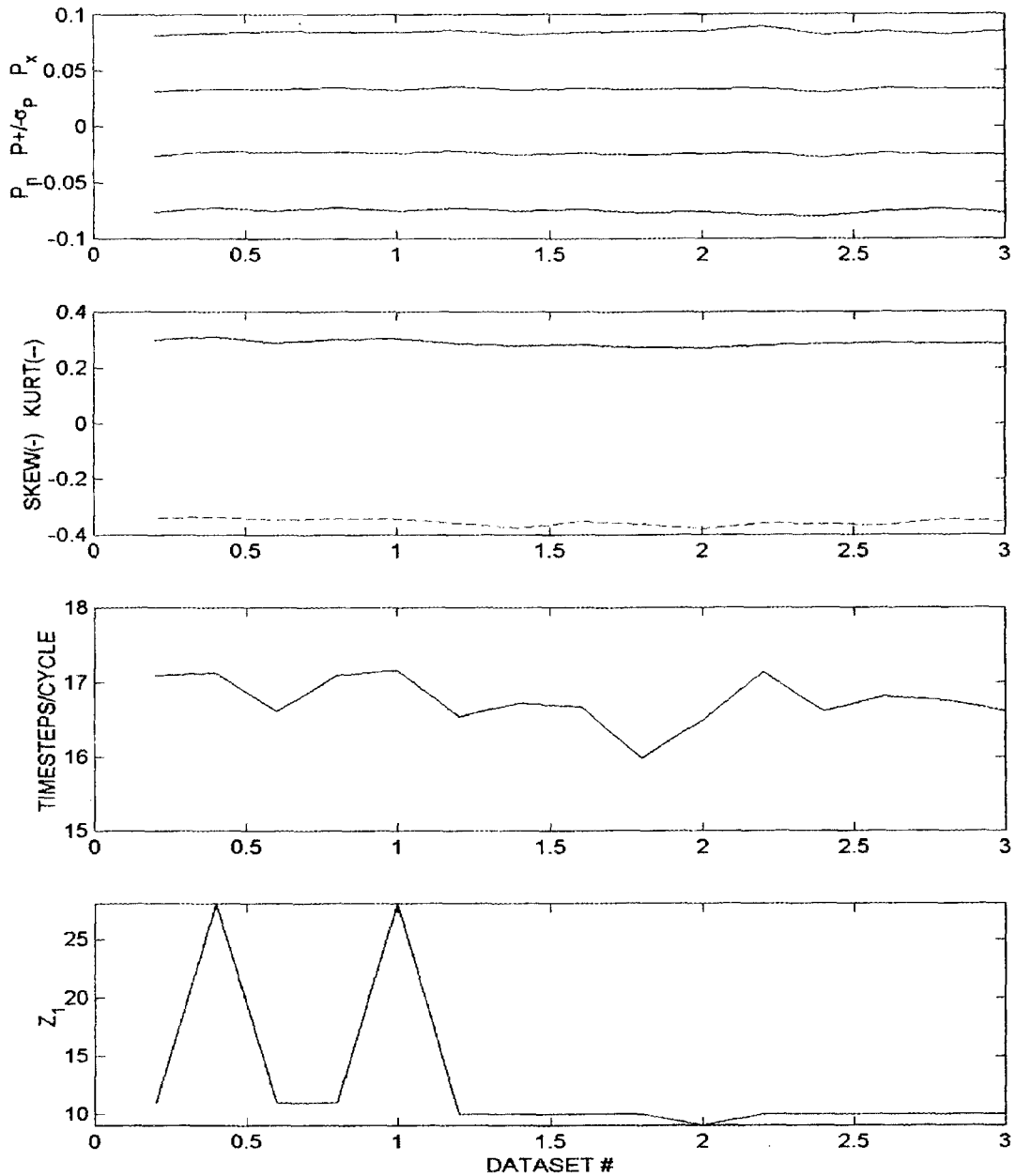



**Fig. 13. Linear measures for the turn-to-turn short seeded-fault.** Top plot shows various linear measures of instantaneous power, $P$: minimum ($P_n$) as the bottom curve, maximum ($P_x$) as the top curve, average plus one standard deviation ($P + \sigma_P$) as the middle top curve, and average minus one standard deviation ($P - \sigma_P$) as the middle bottom curve. Second plot down shows skewness (solid) and kurtosis (- -) in the instantaneous power. Third plot down shows the number of time steps per cycle in the instantaneous power. Bottom plot shows the lag in time steps, corresponding to the first zero in the autocorrelation function. See text for discussion.

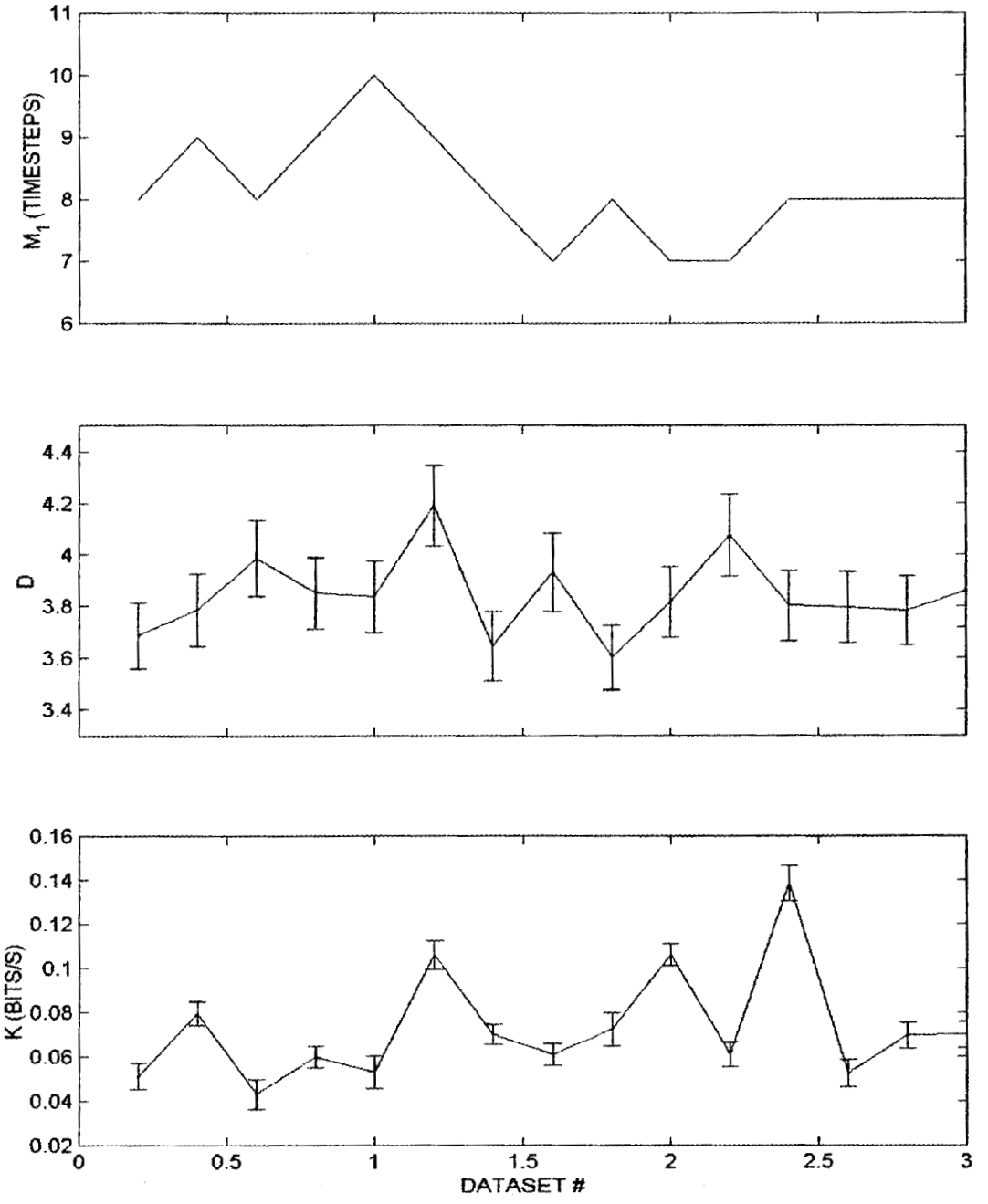



**Fig. 14. Conventional nonlinear measures for turn-to-turn short seeded-fault.** Top plot shows the location (in time steps) of the first minimum in the mutual information function. The middle plot shows the correlation dimension (*D*). The bottom plot shows the Kolmogorov entropy (*K*). Error bars in the middle and bottom plots correspond to the 95% confidence interval. See text for discussion.

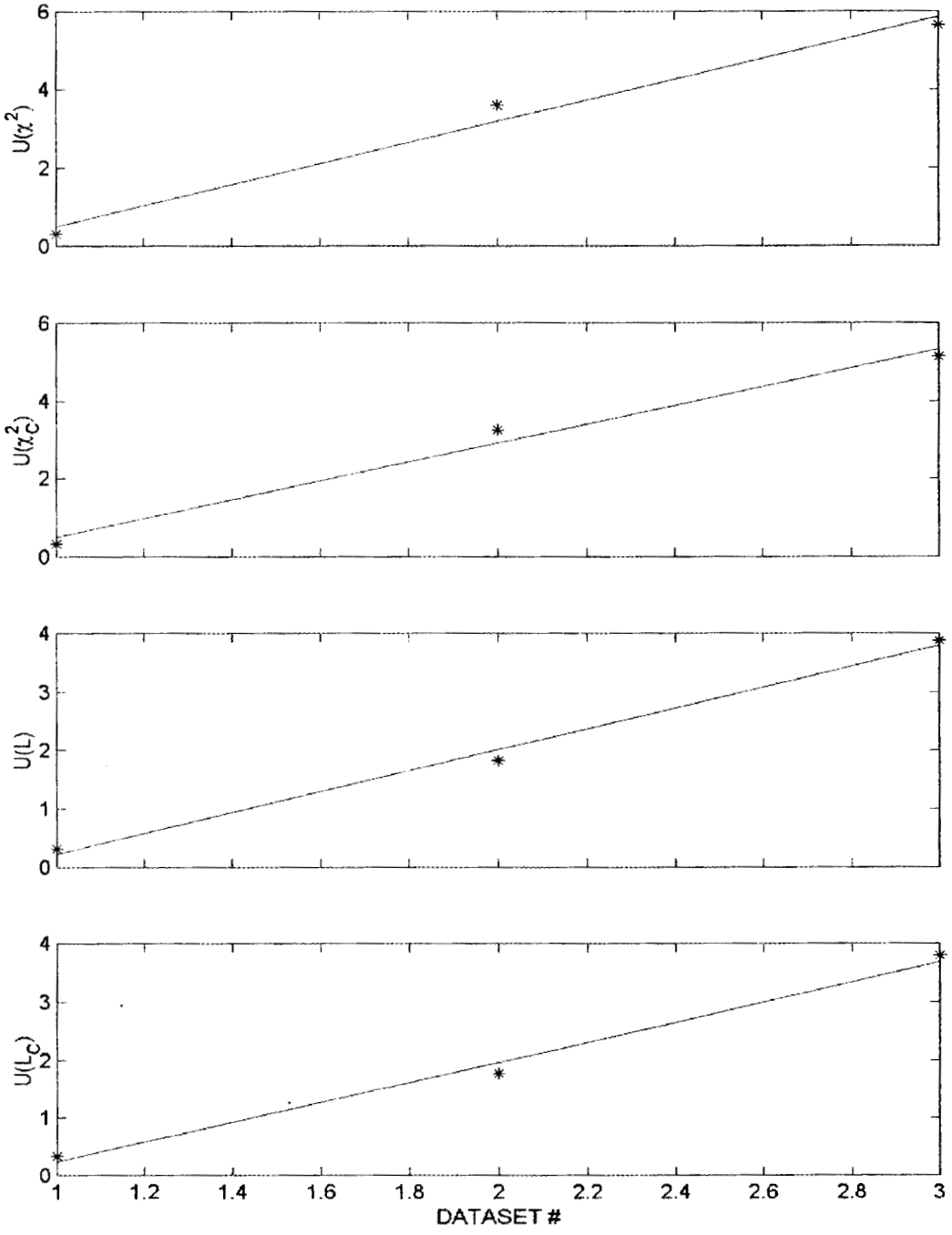



**Fig. 15. PSDM for the turn-to-turn short seeded-fault.** Dataset #1 is for the nominal (no fault) state. Dataset #2 is for the 2.7-ohm short. Dataset #3 is for the 1.35-ohm short. The monotonic rise in the severity of the seeded faults is shown as an almost linear rise (solid line) in the four dissimilarity measures (*) for the chosen set of phase-space parameters.

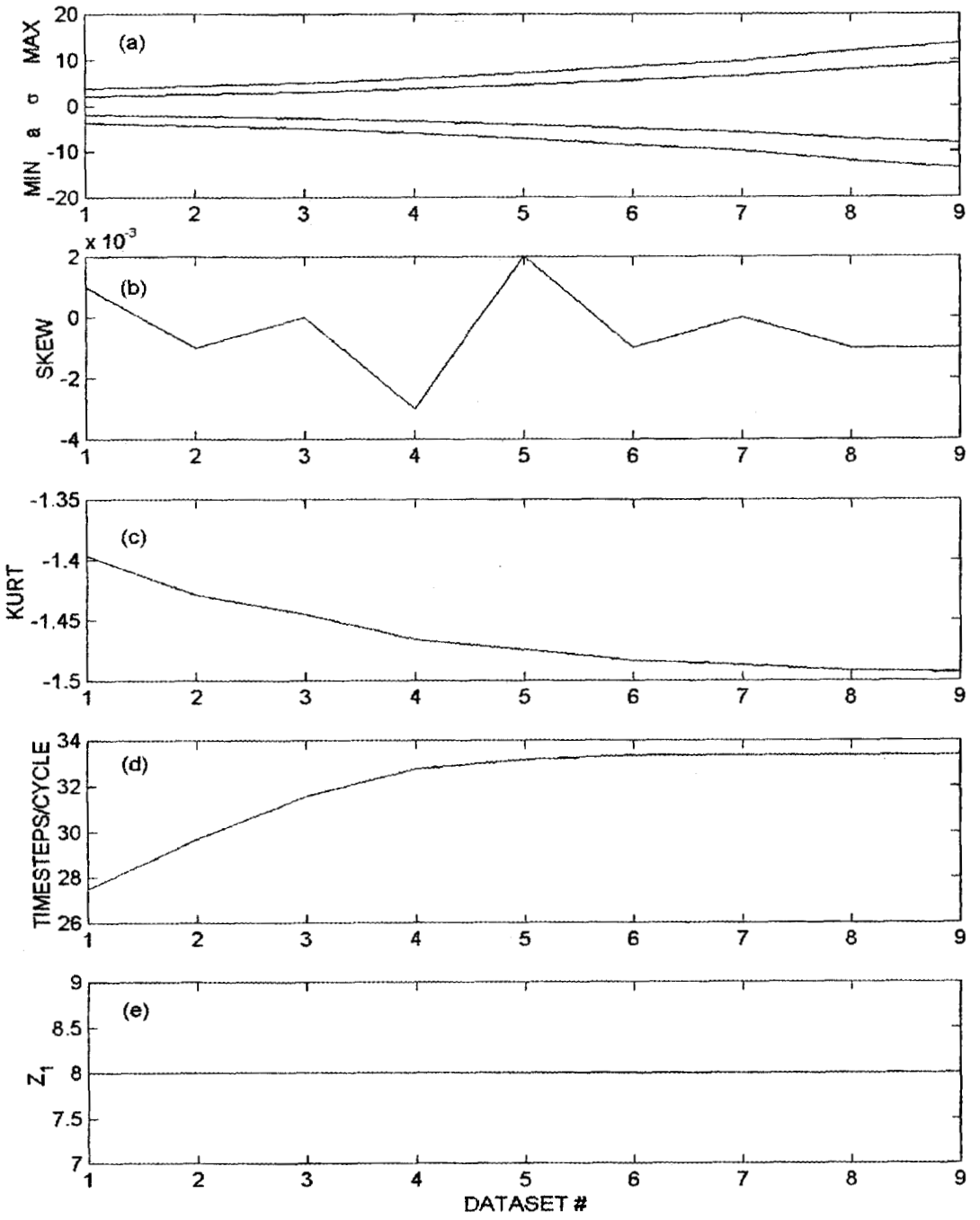



**Fig. 16. Conventional statistical measures vs dataset number.** Results are for the Polytechnic University of Valencia imbalance test sequence: (a) minimum (MIN), maximum (MAX), absolute average deviation ($a$), and sample standard deviation ($\sigma$); (b) skewness; (c) kurtosis; (d) time of time steps per cycle; (e) first zero (in time steps) of auto-correlation function. See text for discussion.

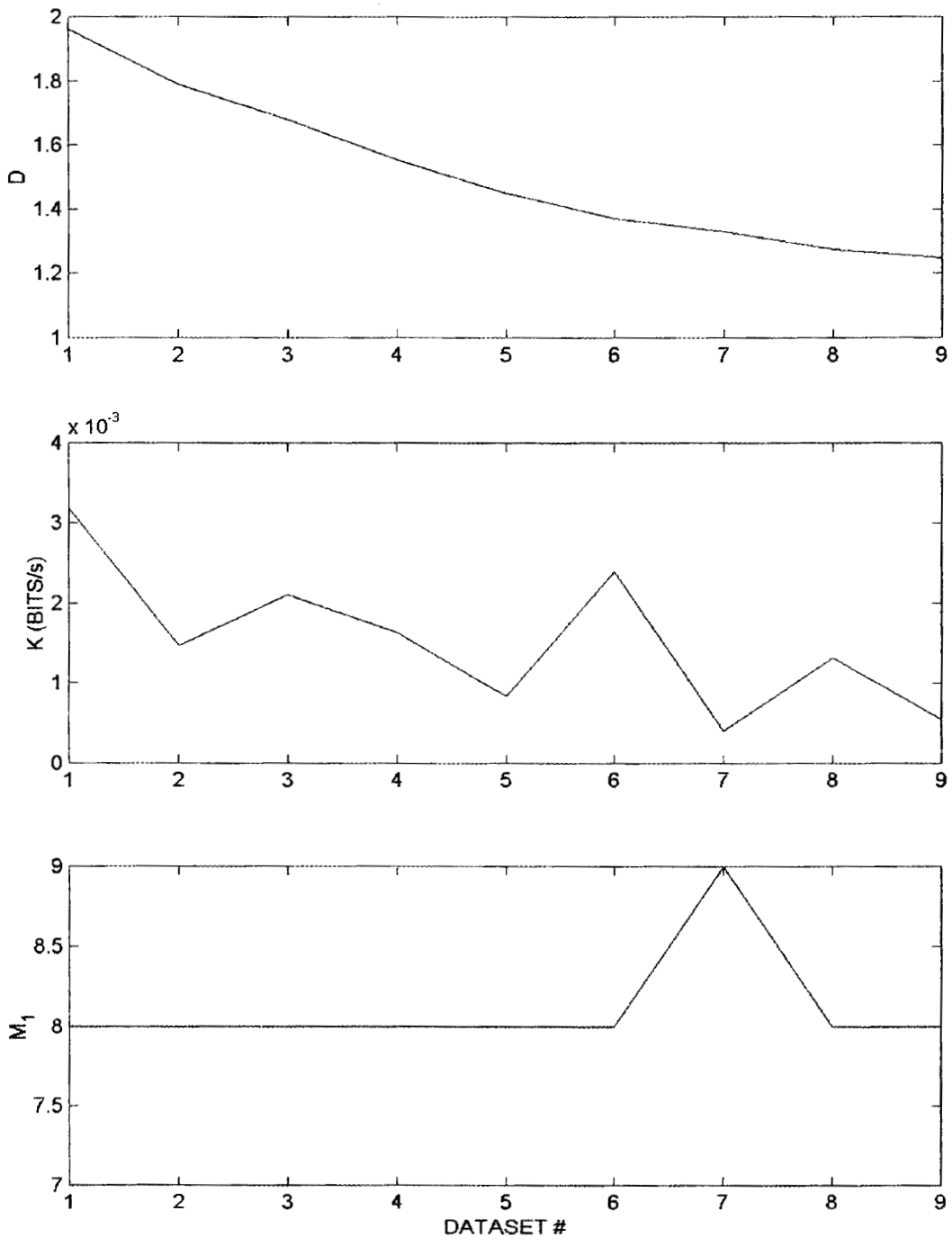



**Fig. 17. Traditional nonlinear measures vs time.** Results are for the Polytechnic University of Valencia imbalance test. (Top) correlation dimension ($D$); (middle) Kolmogorov entropy ($K$); (bottom) location of the first minimum (in time steps) in the mutual information function ($M_1$). See text for discussion.

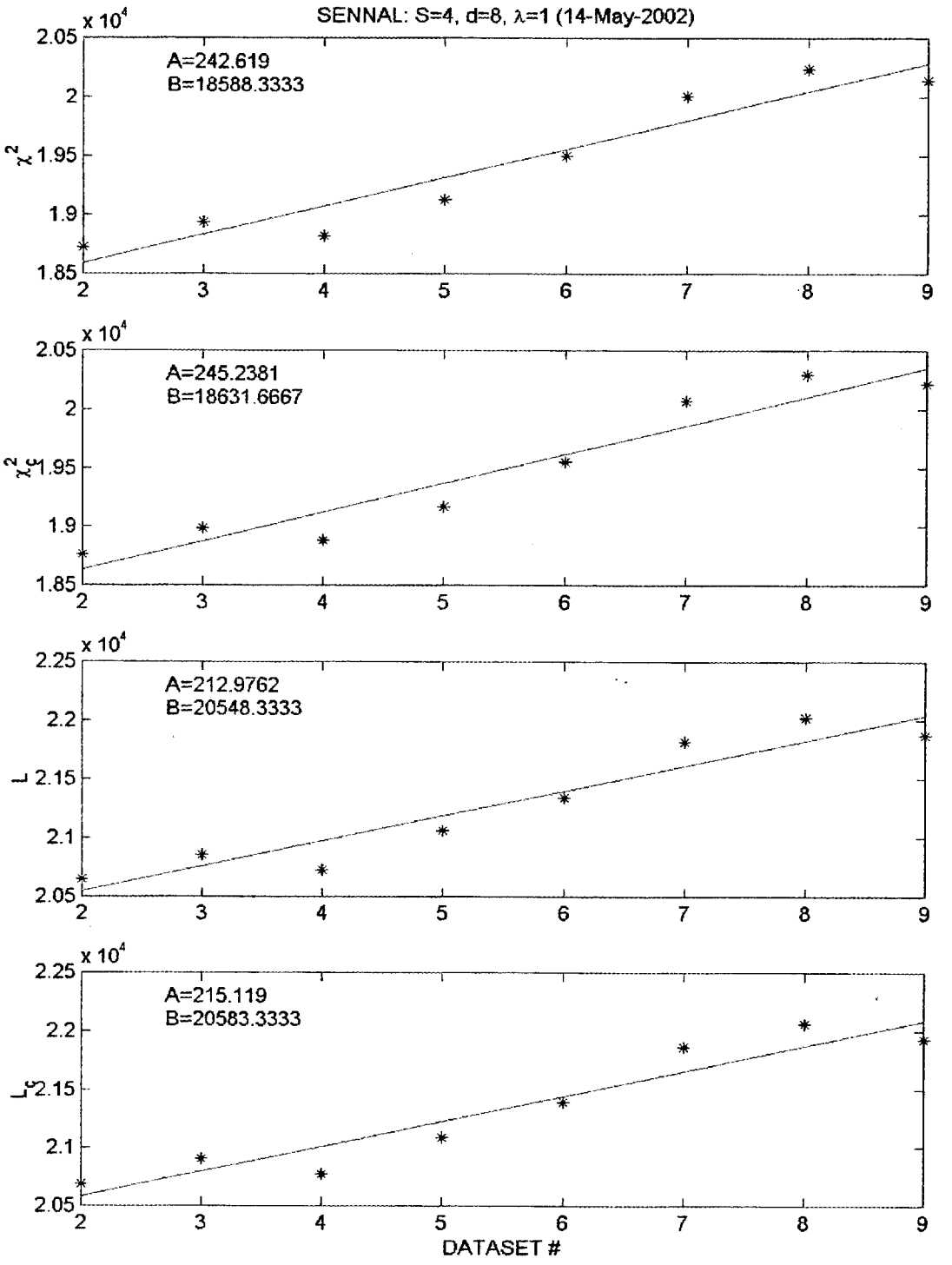
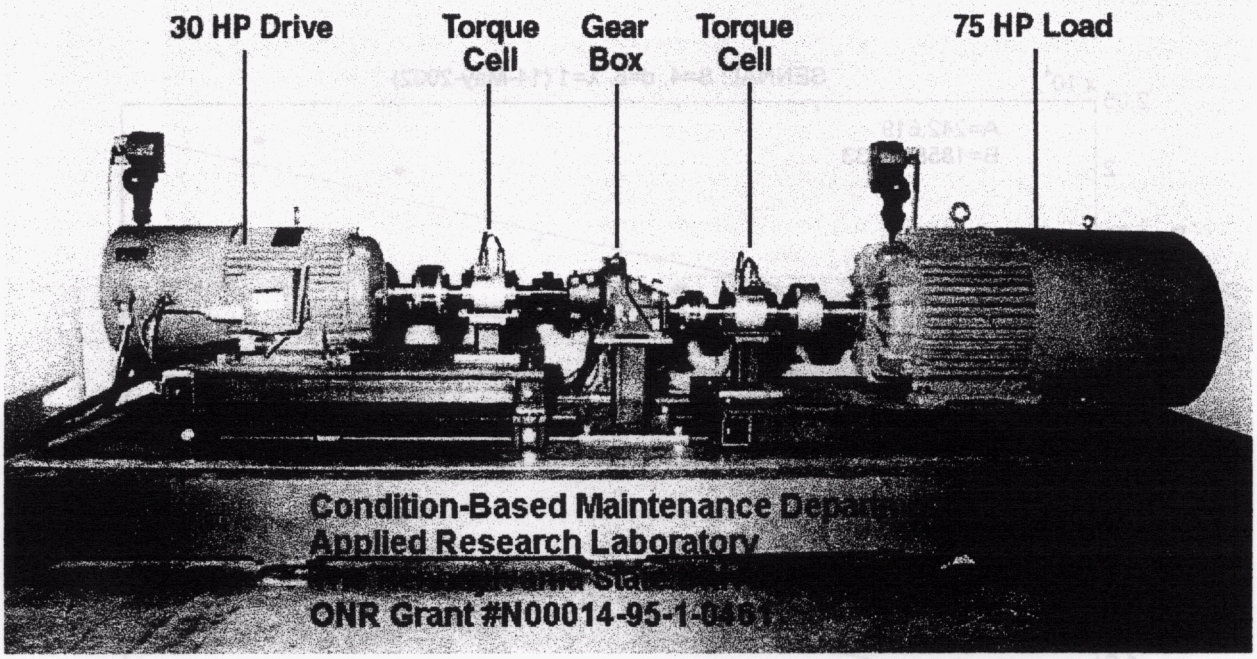



**Fig. 18. Unrenormalized PSDM vs dataset number.** Results are for the Polytechnic University of Valencia imbalance test: (top) $\chi^2$, (second down) $\chi_c^2$, (third down) $L$, (bottom) $L_c$. The values of A and B in each subplot correspond to the slope and y-intercept for the least-squares straight line. See text for discussion.



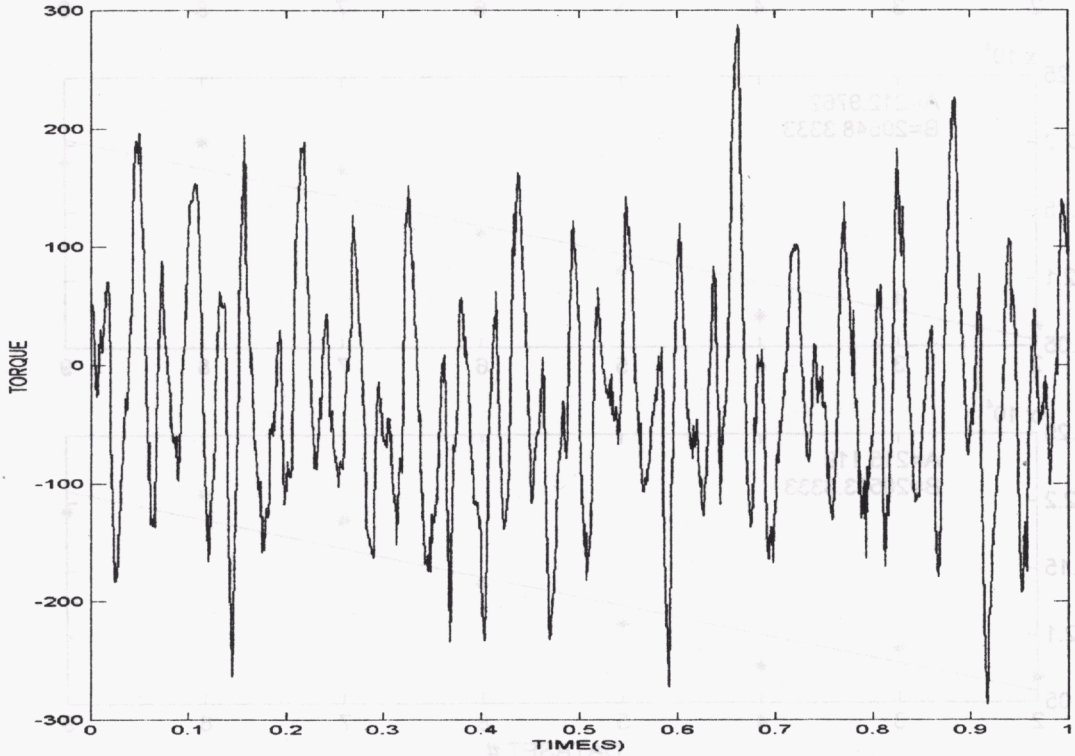




**Fig. 19. Mechanical Diagnostics Test Bed at PSU/ARL.** Top picture shows (from left to right) the 30 hp drive motor, torque cell, gear box, second torque cell, and 75HP electrical generator (load). Lower plot shows a typical segment of data for load torque (in arbitrary units) vs time from the MDTB, sampled at 1 kHz.

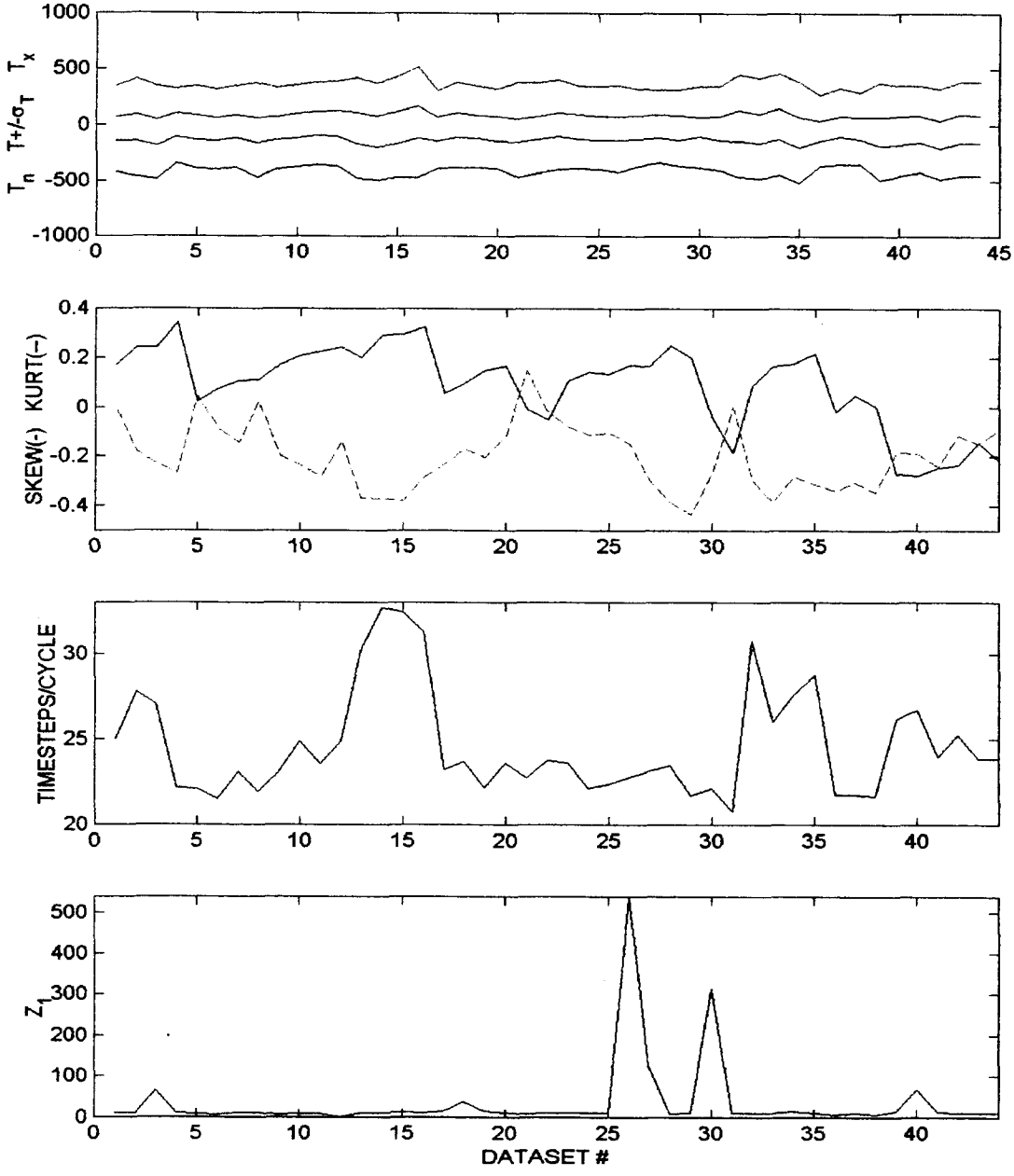



**Fig. 20. Linear measures of load torque from PSU/ARL MDTB.** Top plot shows various linear measures of instantaneous torque, $T$: minimum ($T_n$) as the bottom curve, maximum ($T_x$) as the top curve, average plus one standard deviation ($T + \sigma_T$) as the middle top curve, and average minus one standard deviation ($T - \sigma_T$) as the middle bottom curve. Second plot down shows skewness (solid) and kurtosis (- -) in the instantaneous torque. Third plot down shows the number of time steps per cycle in the torque. Bottom plot shows the lag in time steps, corresponding to the first zero in the autocorrelation function, which varies from zero to 539. See text for discussion.

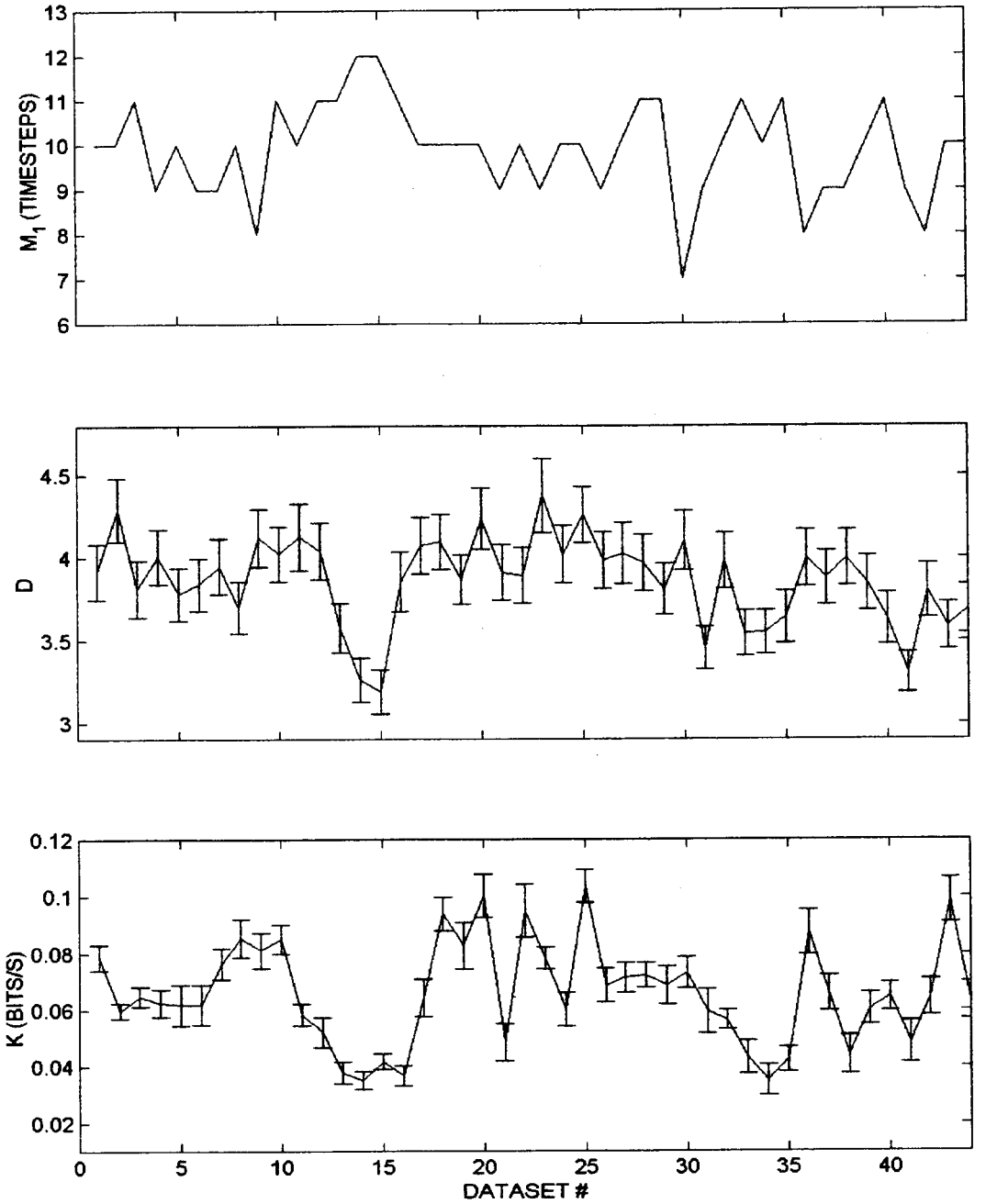



**Fig. 21. Conventional nonlinear measures of load torque data.** Top plot shows the location (in time steps) of the first minimum in the mutual information function. The middle plot shows the correlation dimension (*D*). The bottom plot shows the Kolmogorov entropy (*K*). Error bars in the middle and bottom plots correspond to the 95% confidence interval. See text for discussion.

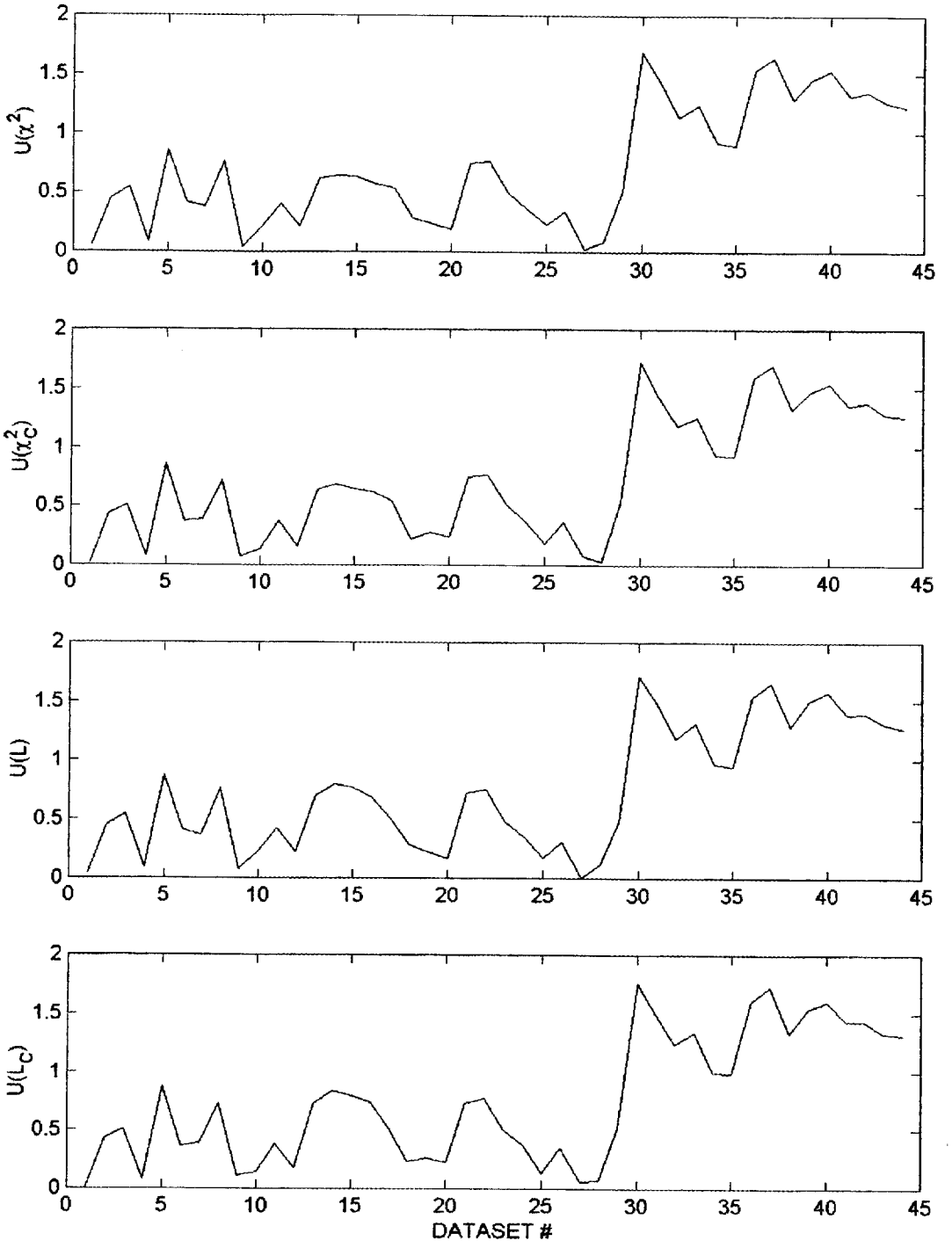



**Fig. 22. PSDM vs dataset number.** Results are for the PSU/ARL MDTB test sequence vs dataset number. Datasets #1–5 are the basecases. These results are for the following phase-space reconstruction parameters: $d$=16, $S$=3, and $\lambda$=61. See text for discussion.

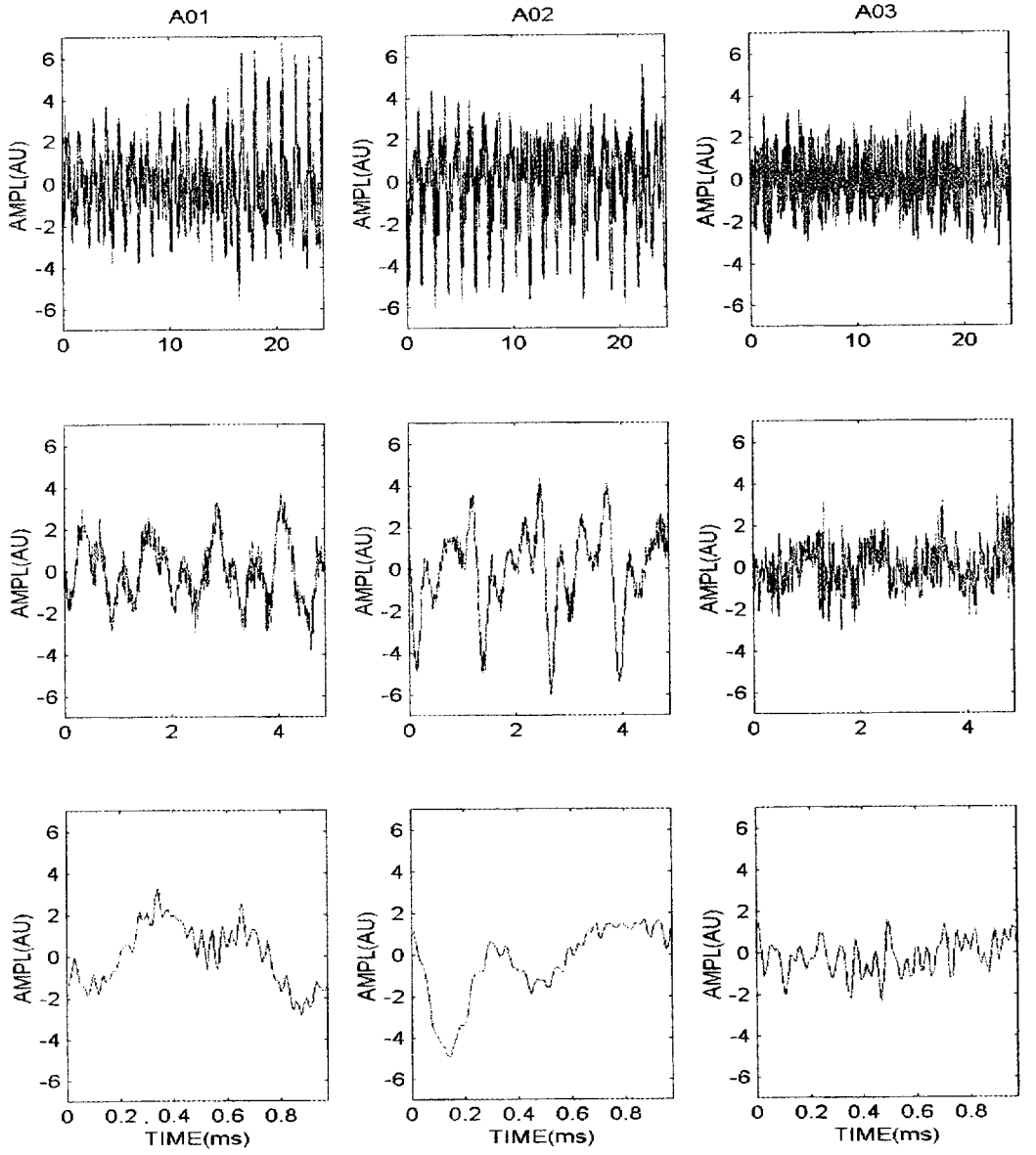



**Fig. 23. Raw tri-axial accelerometer data.** The subplots are for RUN33 of the PSU MDTB test sequence vs time: left column ($A_1$), center column ($A_2$), and right column ($A_3$). Each row (down) of subplots corresponds to five-fold greater resolution in time.

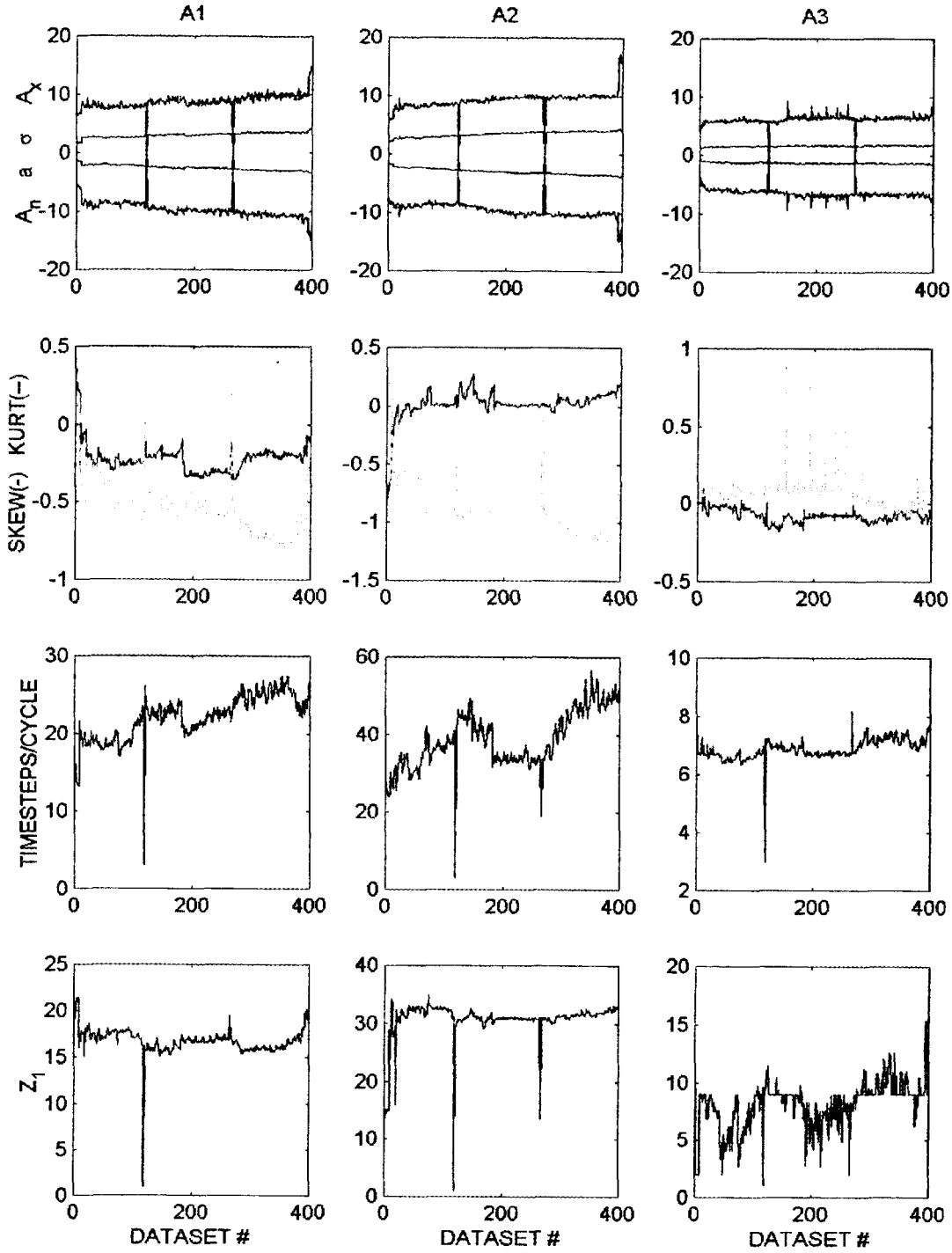



**Fig. 24. Conventional statistical measures of tri-axial acceleration.** The subplots are as follows: (top row) minimum ($A_n$), absolute average deviation ($a$), sample standard deviation ($\sigma$), and maximum ($A_x$); (second row from top) skewness (-) and kurtosis (- -); (third row from top) number of time steps per cycle; (bottom row) first zero (in time steps) in the autocorrelation function.

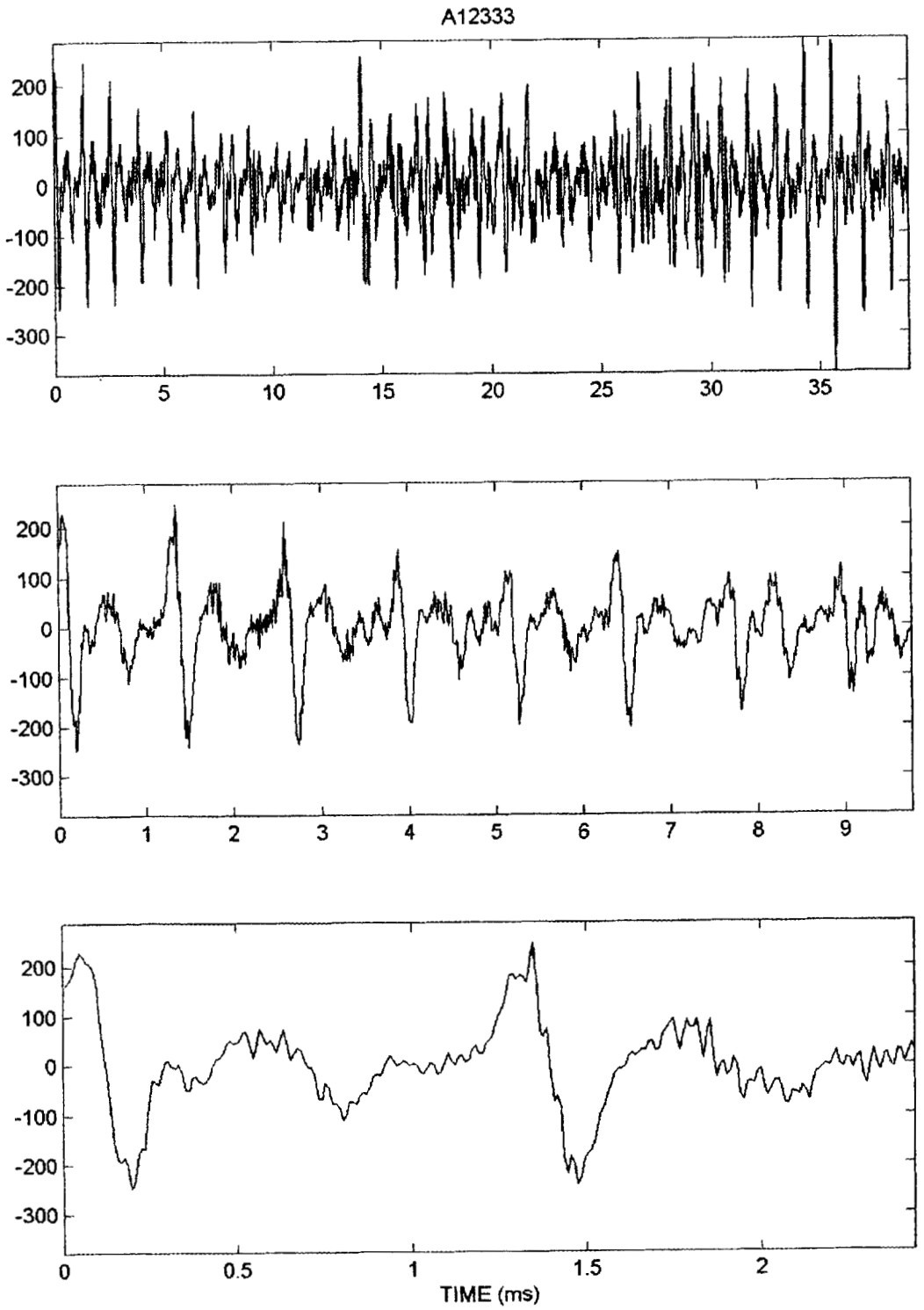



**Fig. 25. Tri-axial accelerometer power vs time.** Subplots are at three successively faster time scales. See text for discussion.

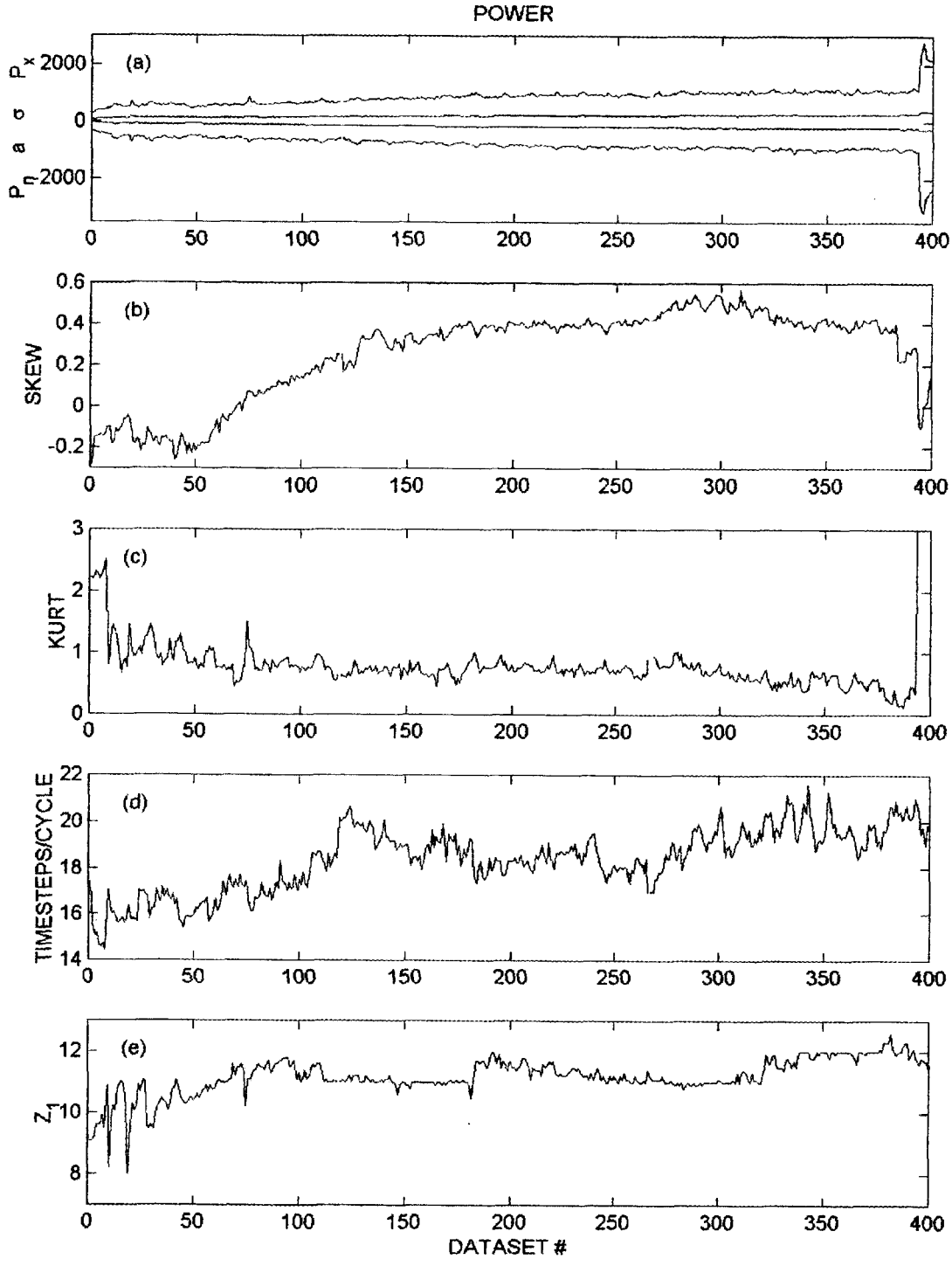



**Fig. 26. Conventional statistical measures of the accelerometer power.** Subplots are for RUN33 of the PSU MDTB test sequence are as follows: (a) minimum ($P_n$), maximum ($P_x$), absolute average deviation ($a$), and standard deviation ($\sigma$); (b) skewness; (c) kurtosis; (d) number of time steps per cycle; (e) first zero ($Z_1$) in the autocorrelation function. See text for discussion.

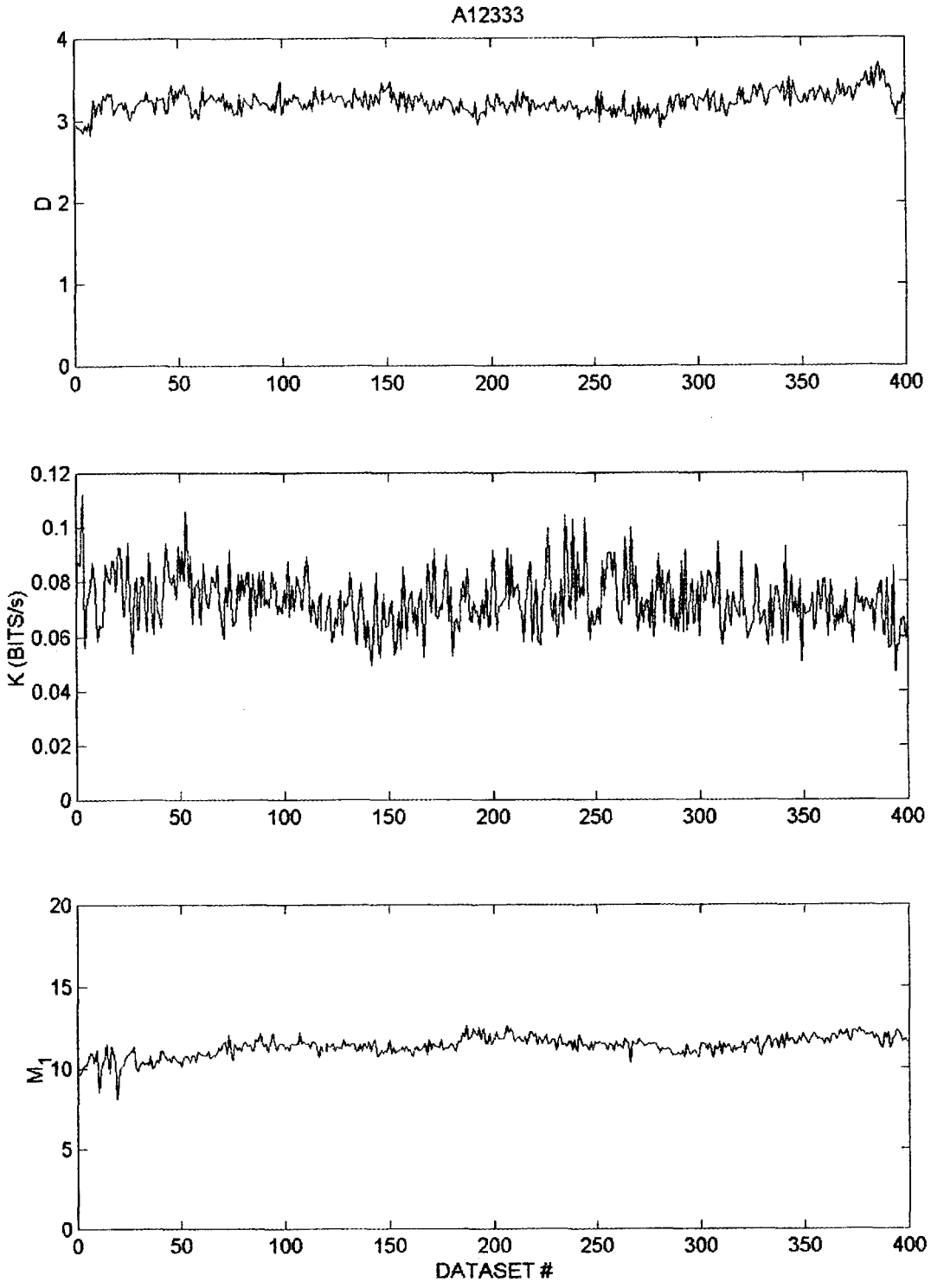



**Fig. 27. Conventional nonlinear measures vs time.** Subplots for RUN33 of PSU MDTB test sequence are as follows: (top) correlation dimension ($D$); (middle) Kolmogorov entropy ($K$); (bottom) location (in time steps) of the first minimum in the mutual information function ($M_1$). See text for discussion.

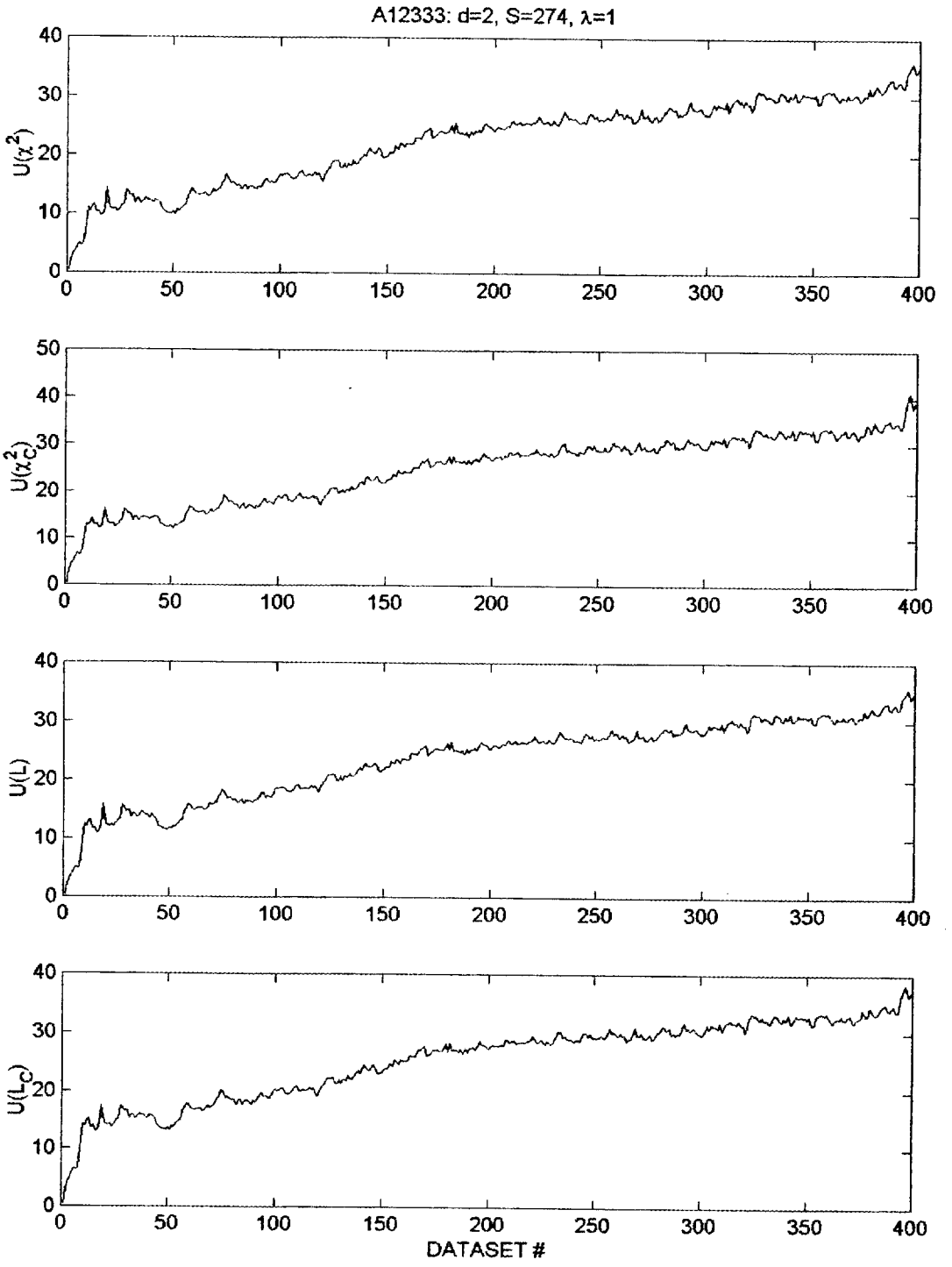



**Fig. 28. PSDM vs dataset number.** Results are for RUN33 of the PSU MDTB gearbox failure sequence for $d$=2, $S$=274, and $\lambda$=1. Results for the two bad datasets (#119 and 266) have been replaced by locally averaged values. See text for discussion.

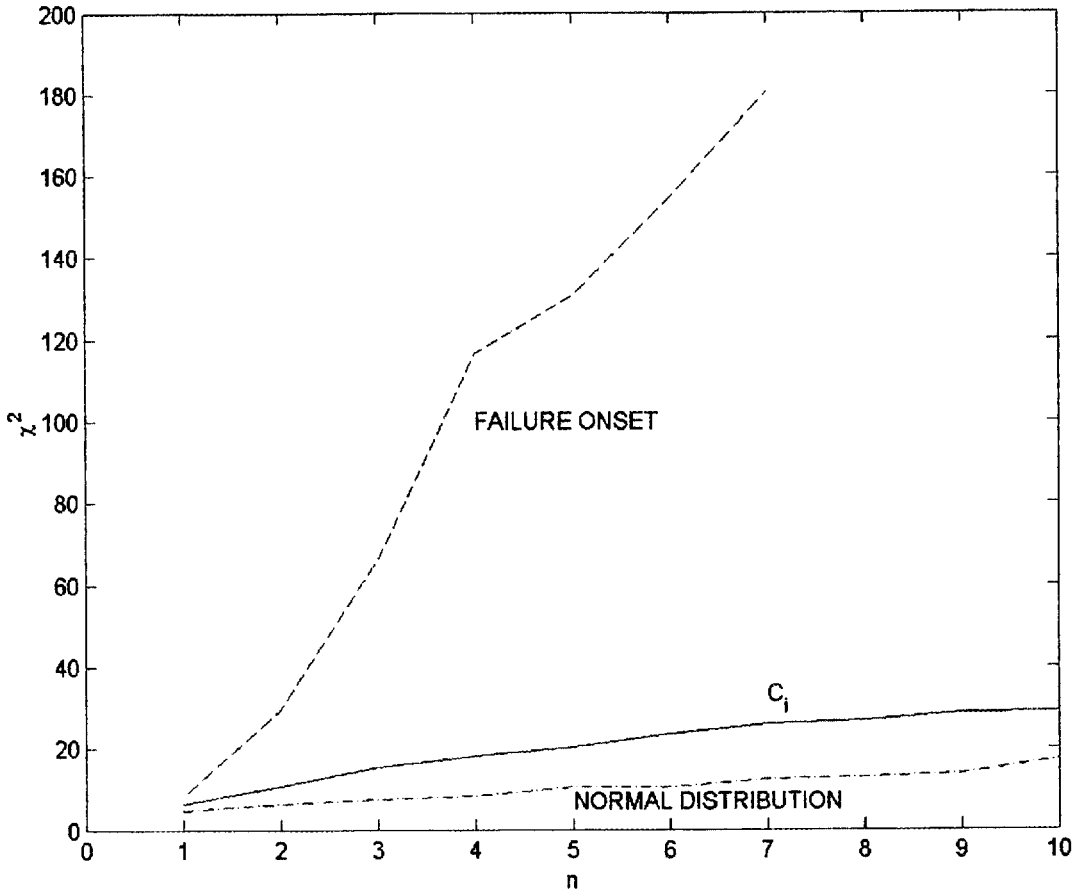



**Fig. 29. Maximum $\chi^2$ vs number ($n$) of sequential points.** Results are obtained for samples from (bottom curve) a normal distribution with zero mean and unity sample standard deviation; (middle curve) composite measure, $C_i$, of condition change from the 200 datasets that span the straight-line fit; (top curve) composite measure, $C_i$, of condition change during failure onset (datasets #394–400). The middle and top curves use the same analysis parameters as in Fig. 29. See text for discussion.

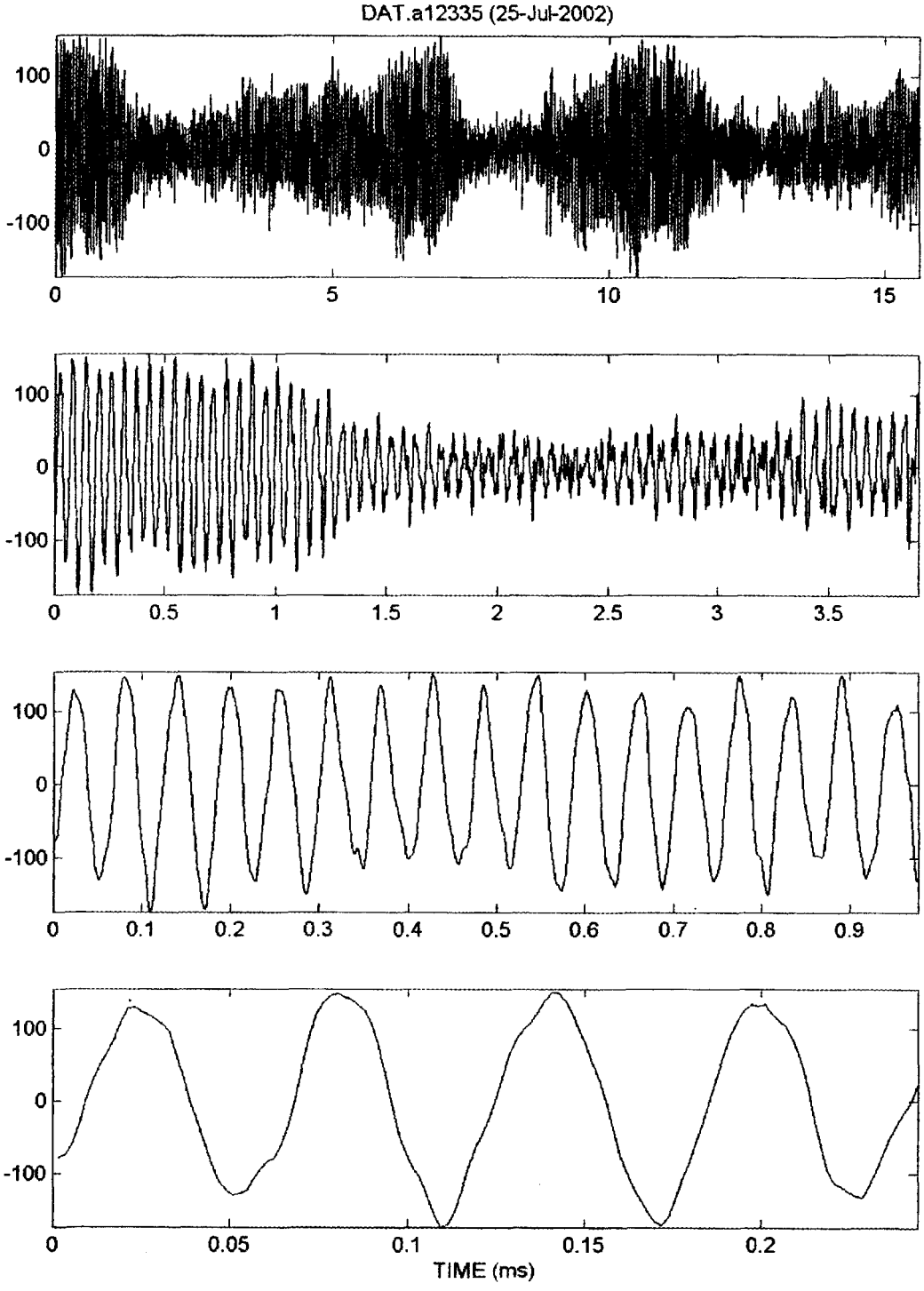



**Fig. 30. Tri-axial accelerometer power vs time.** Results are for RUN35 of the PSU MDTB experiment at four successively faster time scales. See text for discussion.

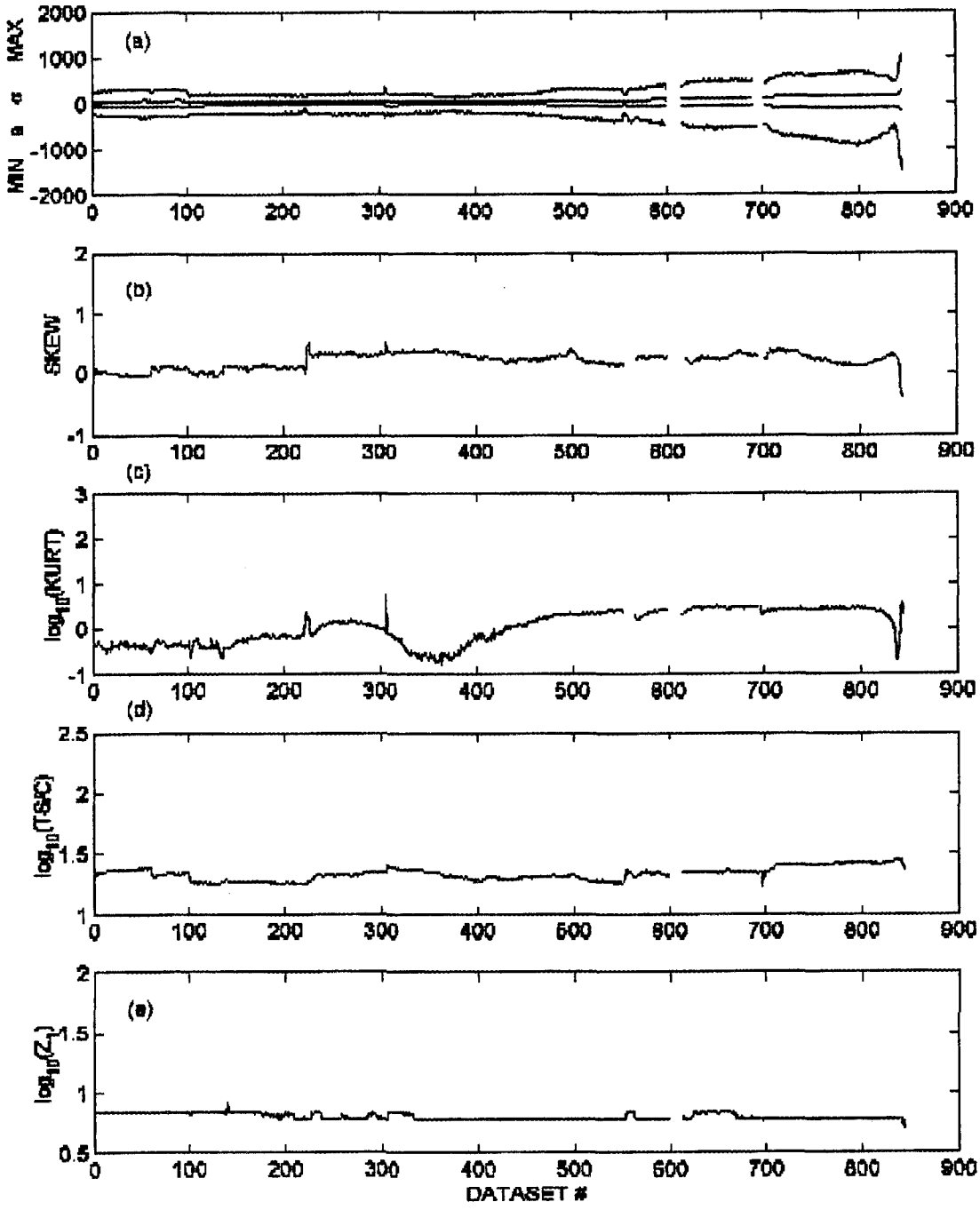



**Fig. 31. Conventional statistical measures of accelerometer power vs dataset number.** Results are for RUN35 of the PSU MDTB test sequence: (a) minimum ($P_n$), maximum ($P_x$), absolute average deviation ($a$), and standard deviation ($\sigma$) of accelerometer power; (b) skewness of time-serial power; (c) kurtosis of power; (d) number of time steps per cycle; (e) first zero ($Z_1$) in the autocorrelation function. See text for discussion.

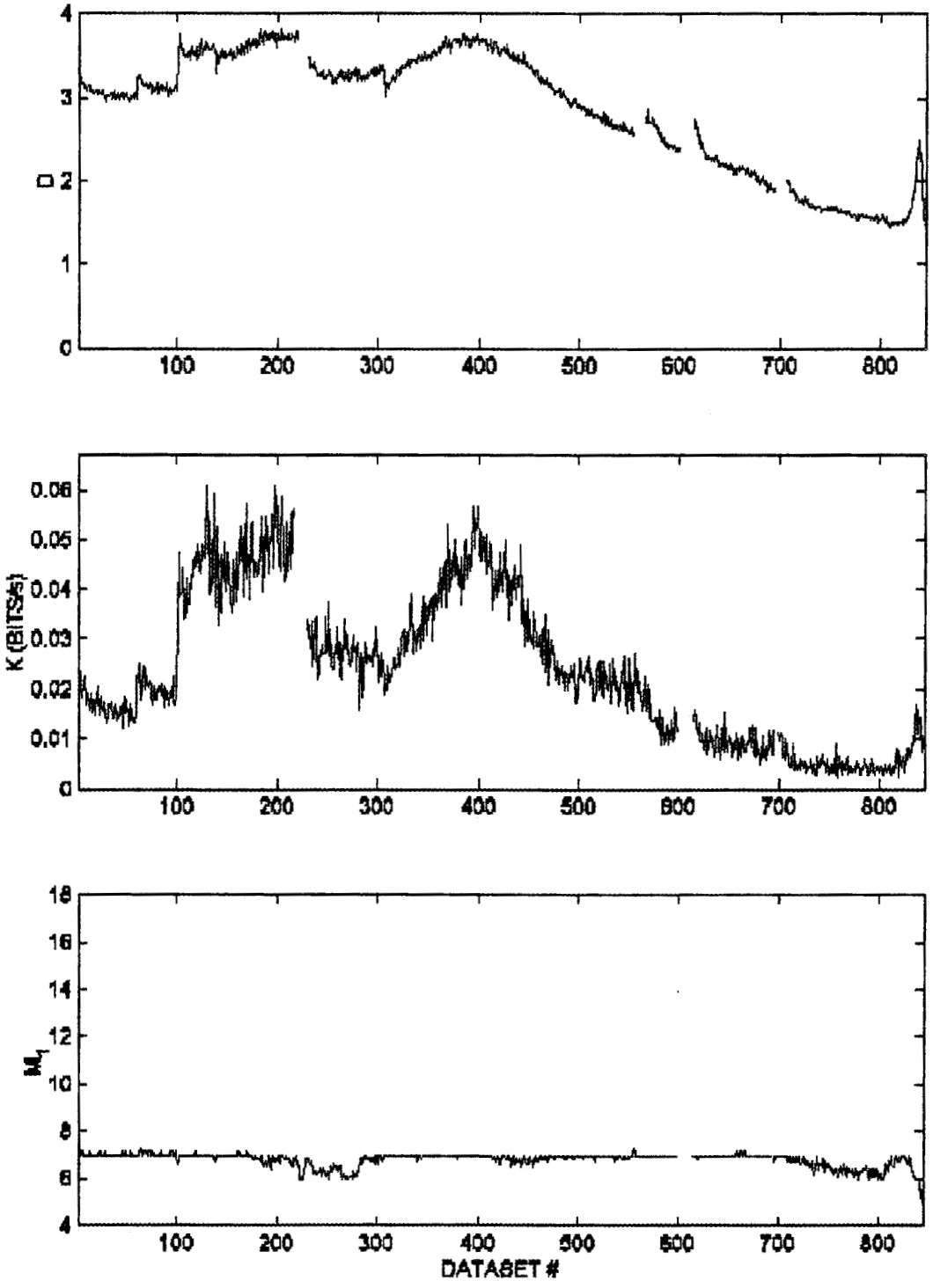



**Fig. 32. Conventional nonlinear measures vs time.** Results are from RUN35 of PSU MDTB test sequence: (top) correlation dimension ($D$); (middle) Kolmogorov entropy ($K$); (bottom) location (in time steps) of the first minimum ($M_1$) in the mutual information function. See text for discussion.

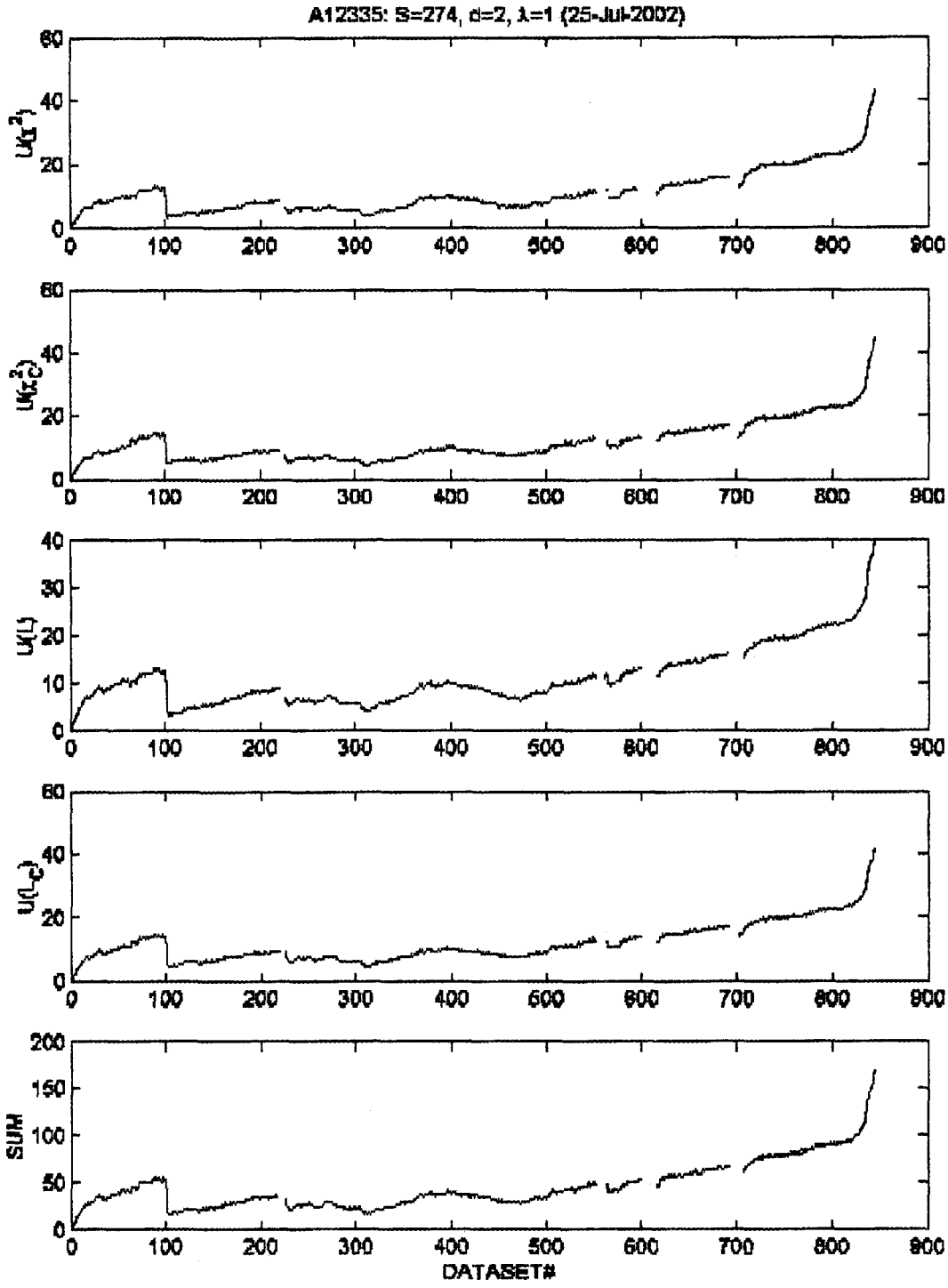



**Fig. 33. PSDM vs dataset number.** Results are for RUN35 of the PSU MDTB gearbox failure sequence for $d$=2, $S$=274, and $\lambda$=1. See text for discussion.

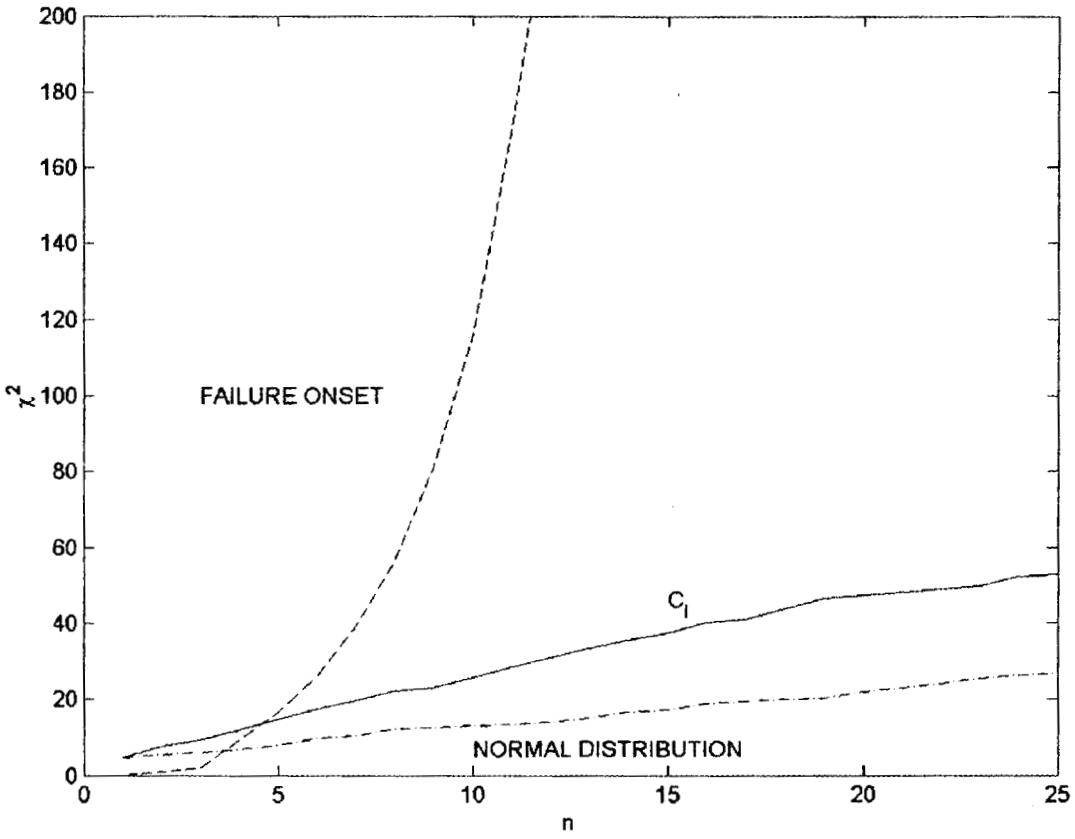



**Fig. 34. Maximum $\chi^2$ vs number ($n$) of sequential points.** Results are obtained from the sample distribution for (bottom curve) a normal distribution with zero mean and unity sample standard deviation; (middle curve) composite measure, $C_i$, of condition change from the 100 datasets that span the straight-line fit; (top curve) composite measure, $C_i$, of condition change during failure onset (datasets #820–831). The middle and top curves use the same analysis parameters as in Fig. 30. See text for discussion.

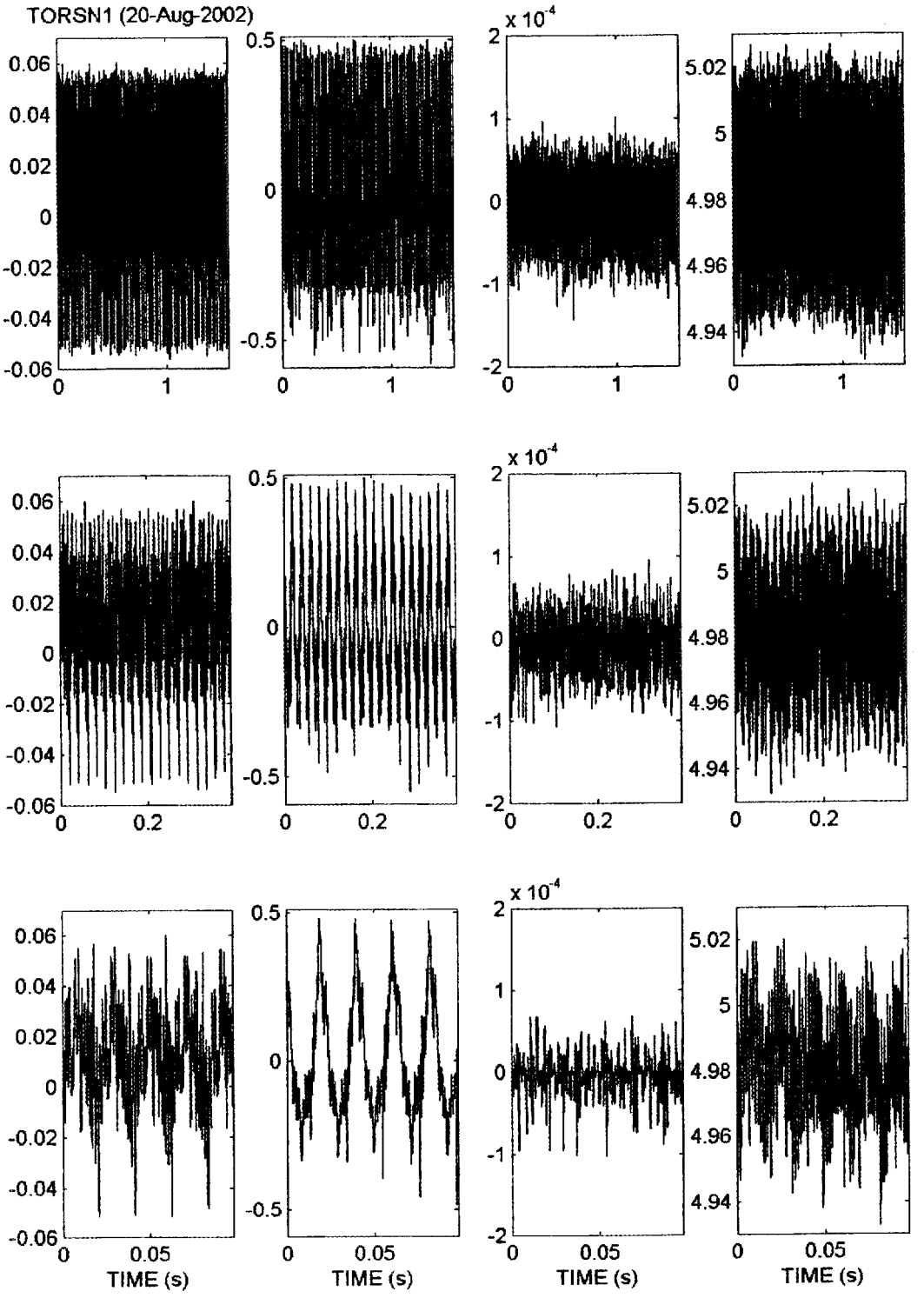



**Fig. 35. Four data channels from PSU TORSION data.** Results are at three successively faster time scales: (left two columns) horizontal and vertical acceleration; (third column from the left) AC-coupled motor power; (right-most column) DC-coupled motor power. See text for discussion.

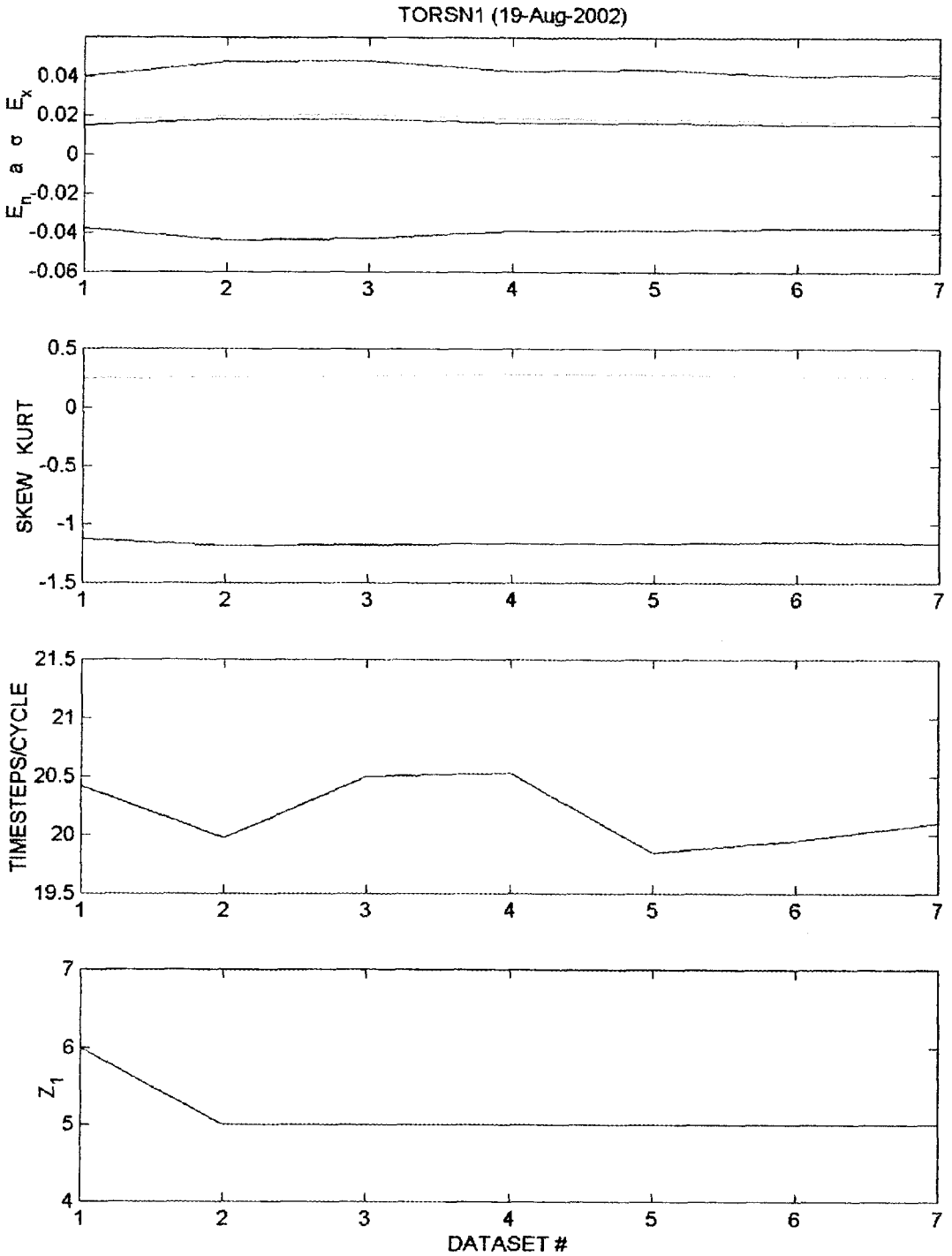



**Fig. 36. Conventional statistical measures of DC motor power vs dataset number.** Results are for the PSU TORSION experiment: (top) minimum ($P_n$), maximum ($P_x$), absolute average deviation ($a$), and standard deviation ($\sigma$); (second from top) skewness and kurtosis of power; (third from top) number of time steps per cycle; (bottom) first zero ($Z_1$) in the autocorrelation function. See text for discussion.

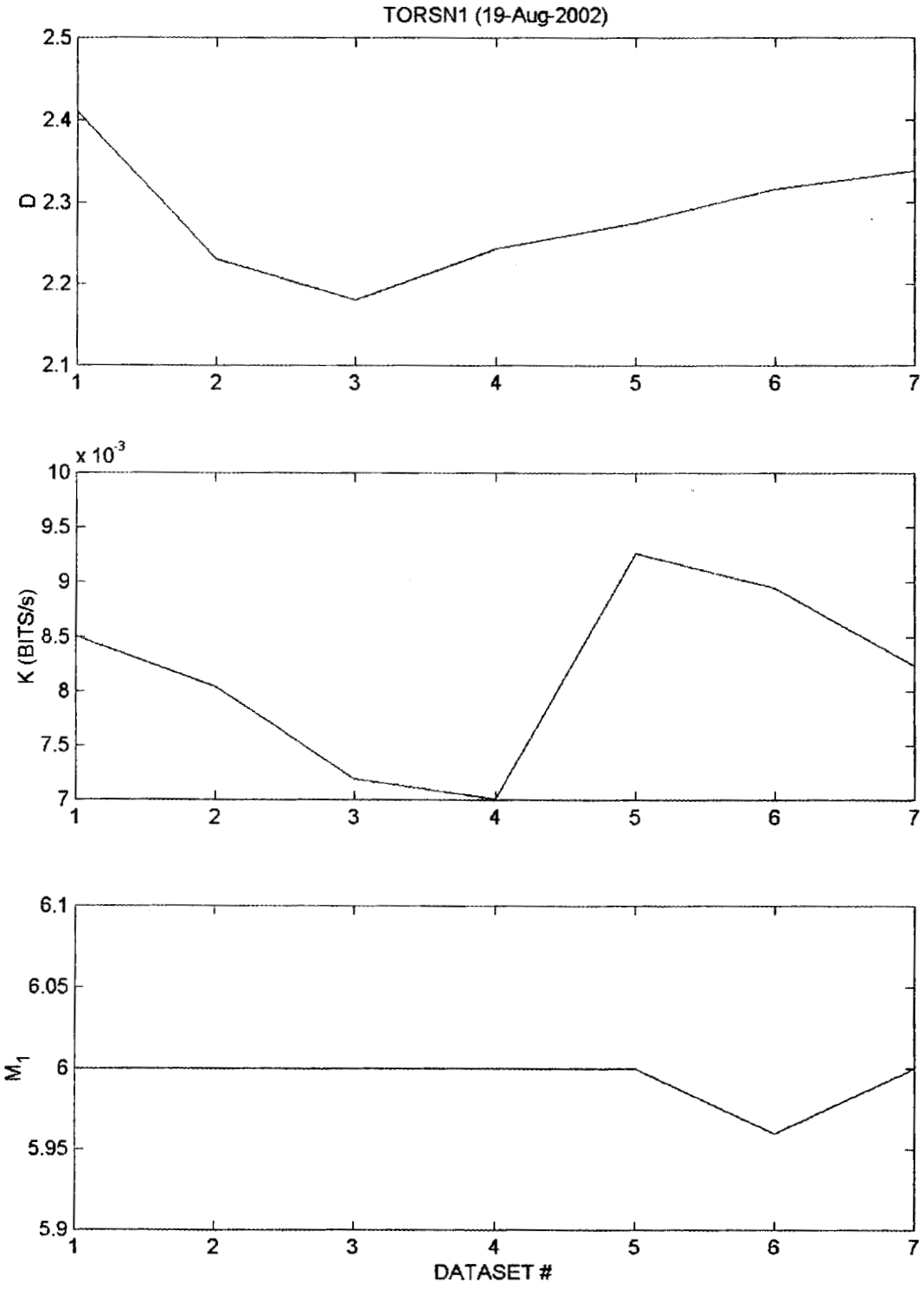



**Fig. 37. Conventional nonlinear measures vs time.** Results are from PSU TORSION experiment: (top) correlation dimension ($D$); (middle) Kolmogorov entropy ($K$); (bottom) location (in time steps) of the first minimum ($M_1$) in the mutual information function. See text for discussion.

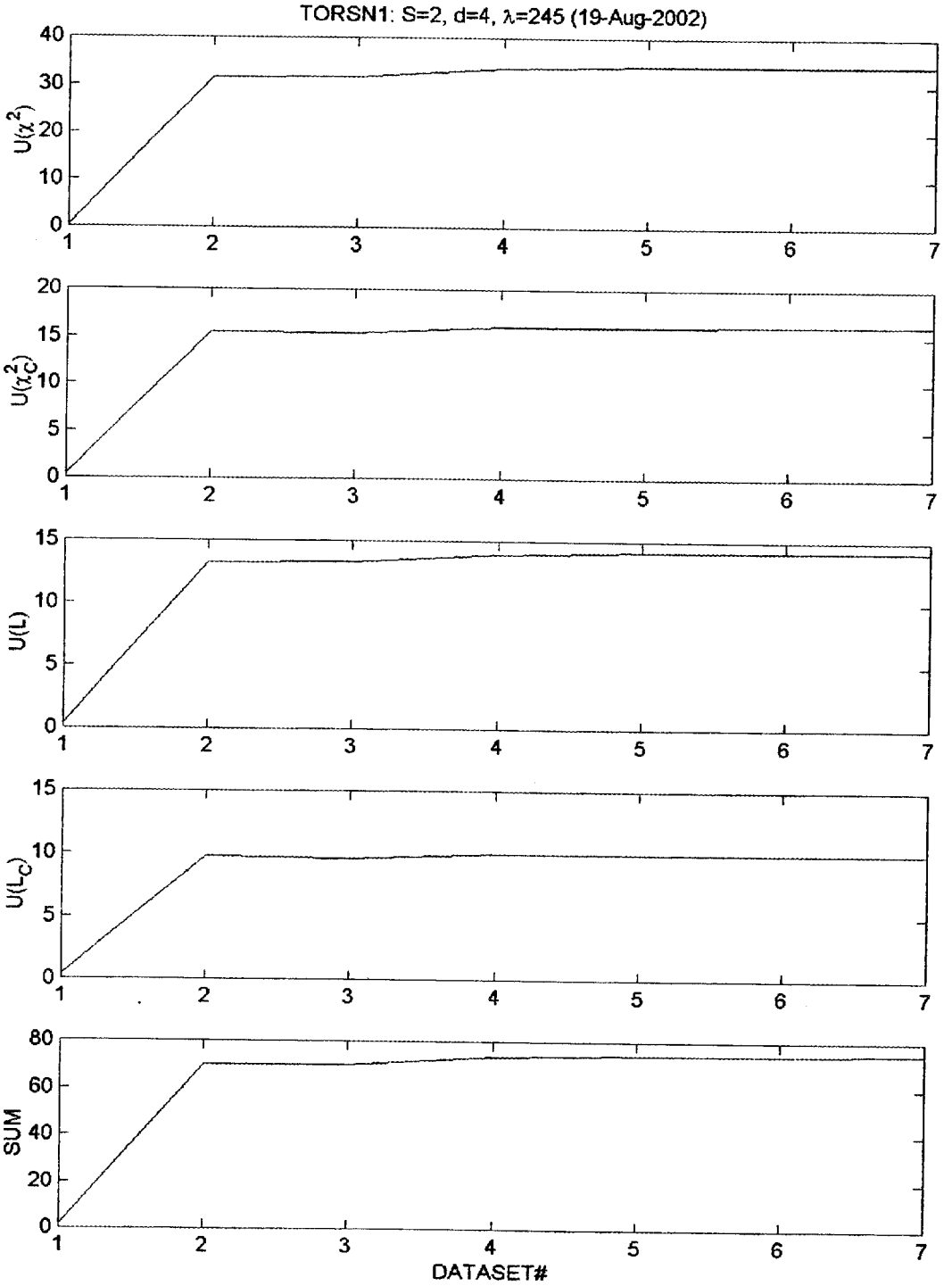



**Fig. 38. PSDM (top four) and their sum (bottom) vs dataset number.** Results are for the test sequence for PSU TORSION experiment for $d$=4, $S$=2, and $\lambda$=245. See text for discussion.

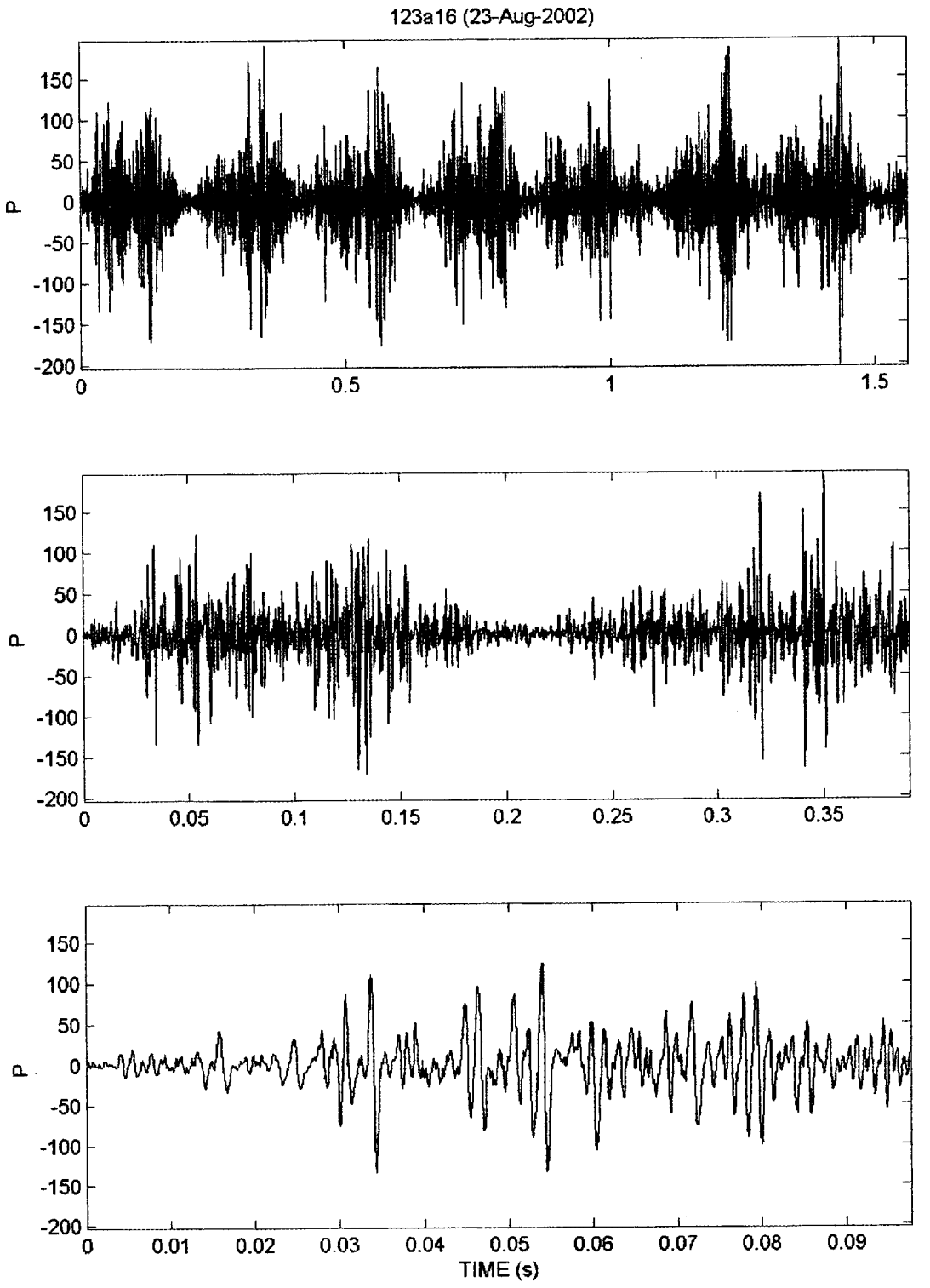



**Fig. 39. Accelerometer power from the PSU BEARING experiment.** Results are shown at three successively faster time scales. See text for discussion.

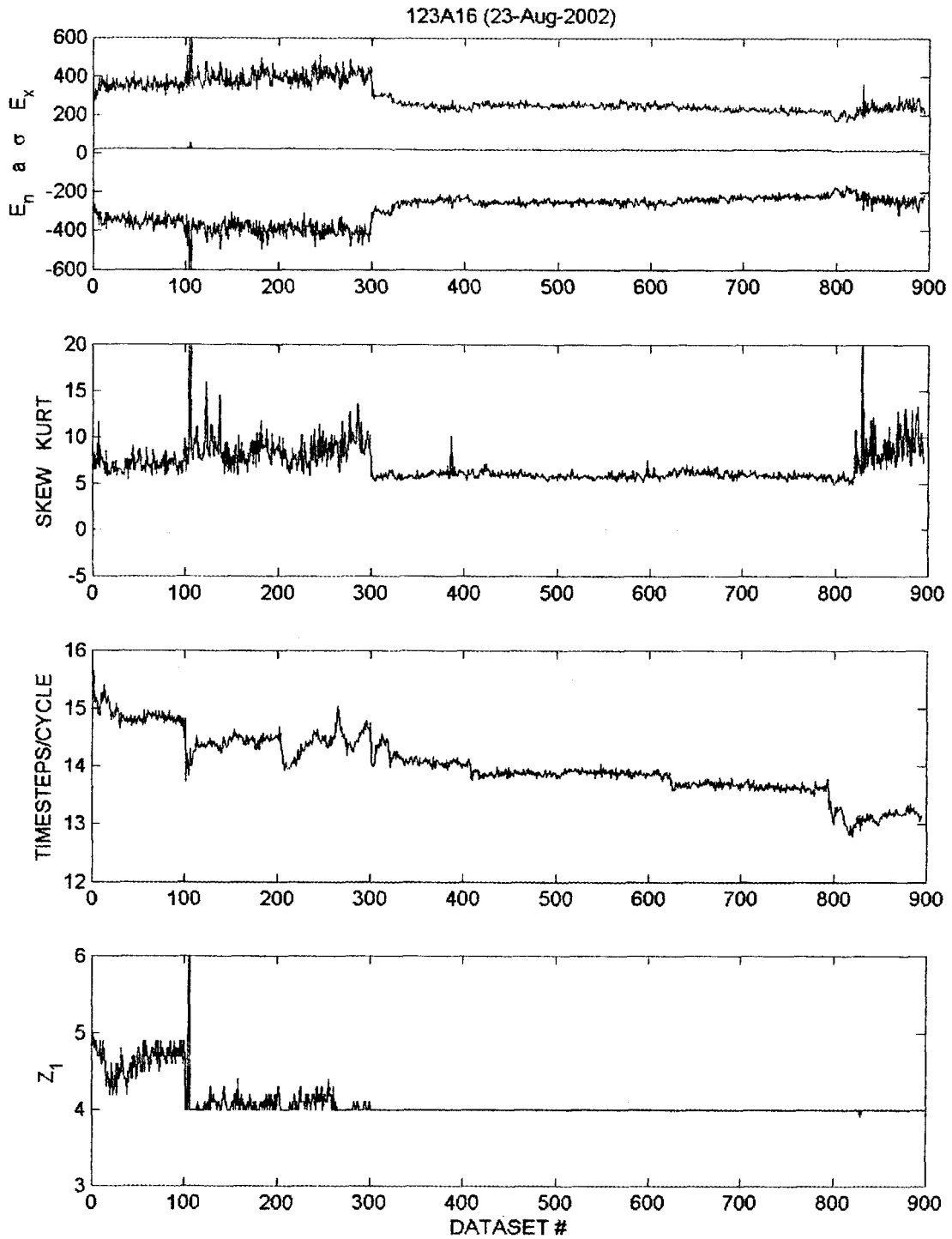



**Fig. 40. Conventional statistical measures of accelerometer power vs dataset number.** Results are for the PSU BEARING experiment: (top) minimum ($P_n$), maximum ($P_x$), absolute average deviation ($a$), and standard deviation ($\sigma$); (second from top) skewness and kurtosis of power; (third from top) number of time steps per cycle; (bottom) first zero ($Z_1$) in the autocorrelation function. See text for discussion.

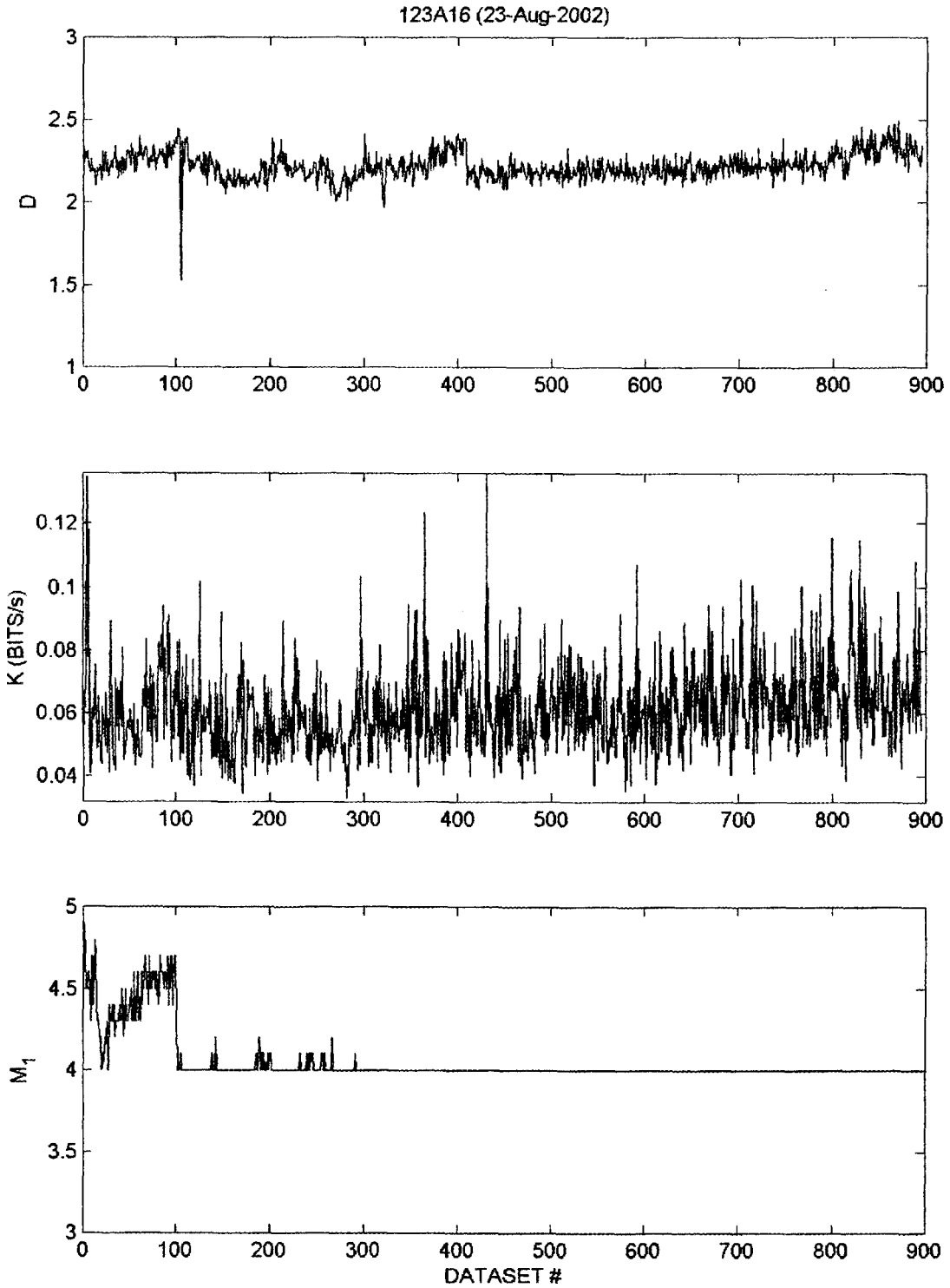



**Fig. 41. Conventional nonlinear measures vs time.** Reslts are for PSU BEARING experiment: (top) correlation dimension ($D$); (middle) Kolmogorov entropy ($K$); (bottom) location (in time steps) of the first minimum ($M_1$) in the mutual information function. See text for discussion.

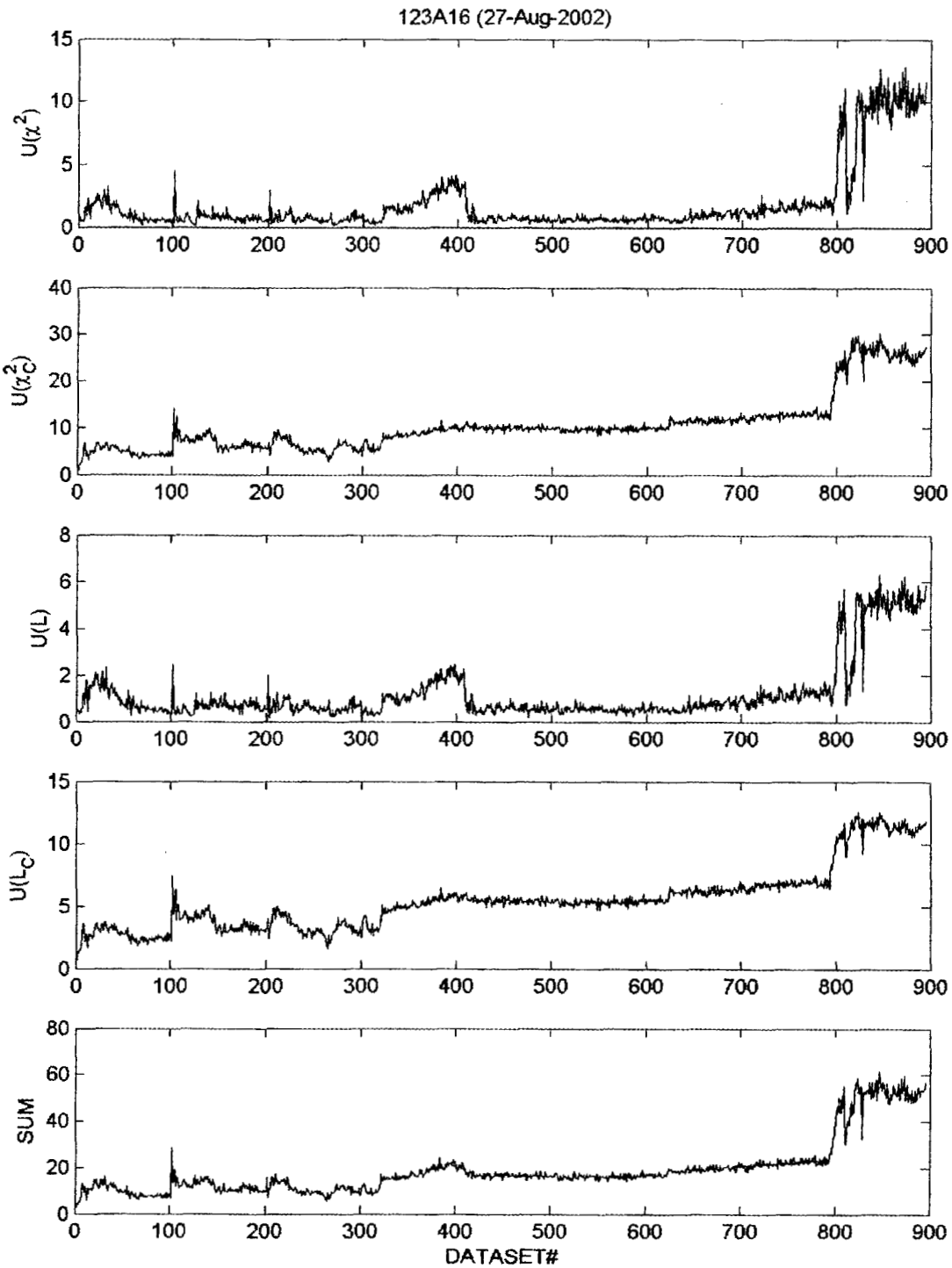



**Fig. 42. PSDM (top four) and their sum (bottom) vs dataset number.** Results are from the test sequence for PSU BEARING experiment for $d$=4, $S$=2, and $\lambda$=73. See text for discussion.

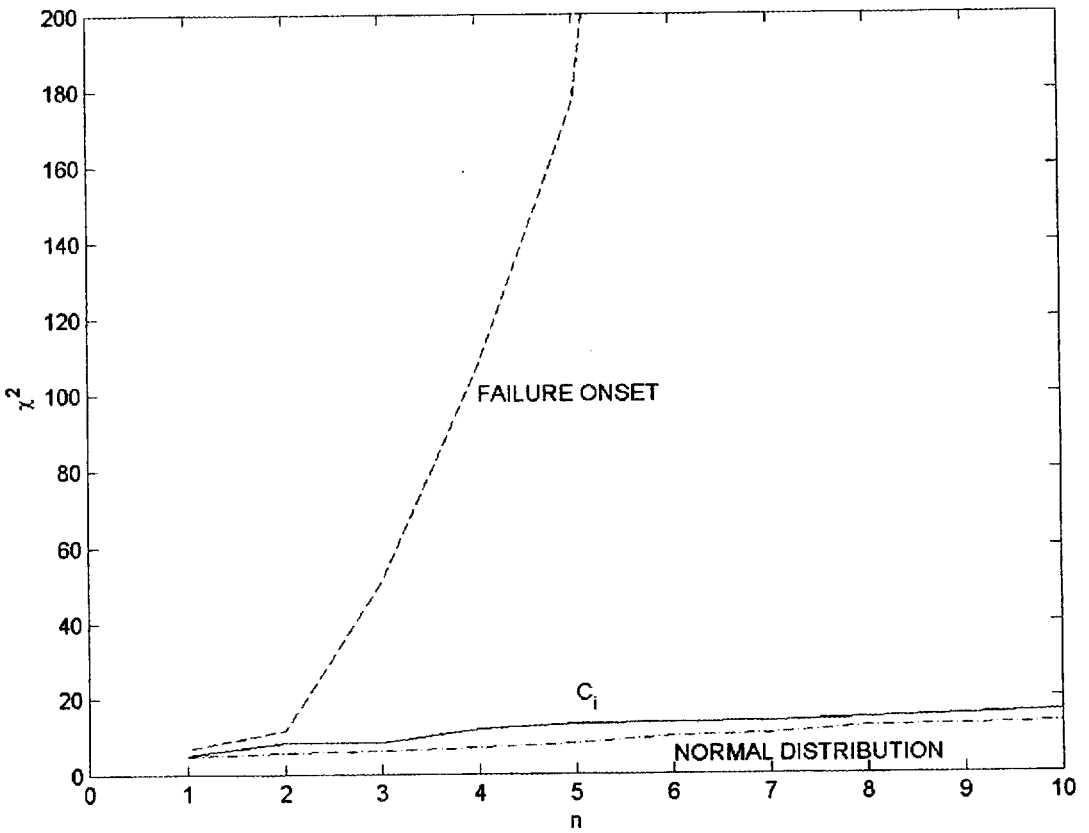



**Fig. 43. Maximum $\chi^2$ vs number ($n$) of sequential points.** Results are obtained from the sample distribution for (bottom curve) a normal distribution with zero mean and unity sample standard deviation; (middle curve) composite measure, $C_i$, of condition change from the 100 datasets that span the straight-line fit; (top curve) composite measure, $C_i$, of condition change during failure onset (datasets #795–804).

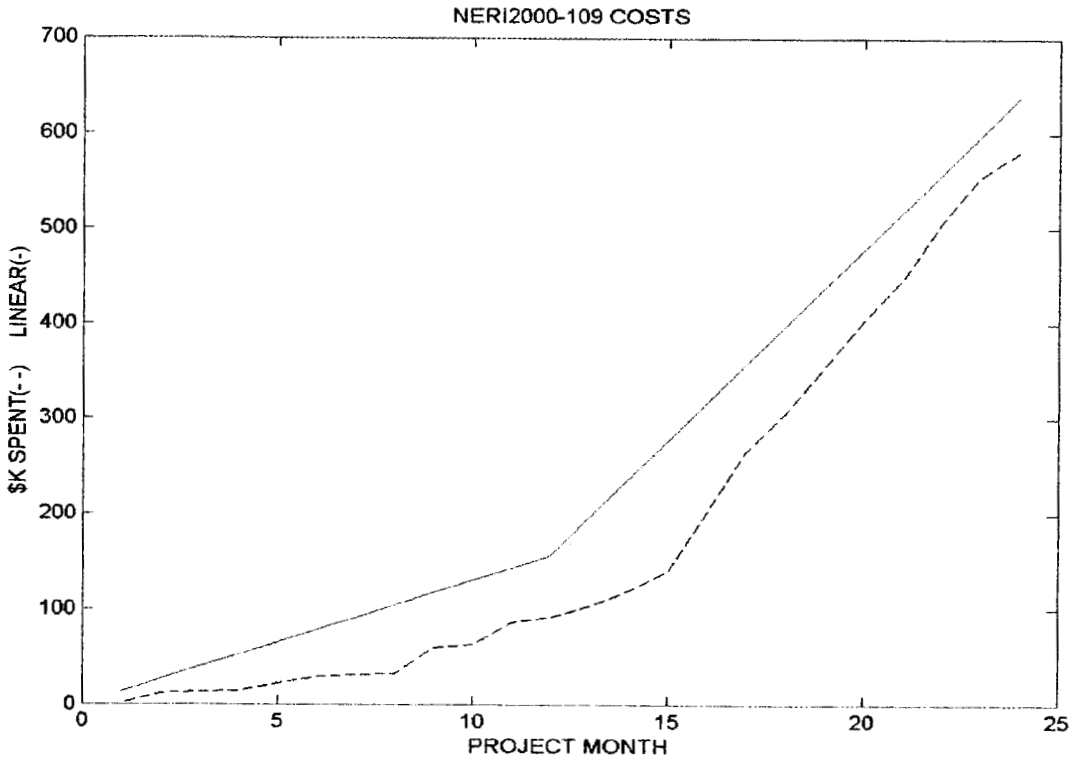



**Fig. 44. Budget and schedule.**

## APPENDIX A: DESCRIPTION OF ANALYSIS METHODS

This Appendix is organized as follows. Section A.1 describes the method for removal of artifacts from the data. Section A.2 discusses conventional statistical measures for time-serial analysis. Section A.3 describes three traditional nonlinear measures for our analysis. Section A.4 explains ORNL's nonlinear PSDM.

### A.1 ARTIFACT REMOVAL

Data frequently include artifacts, such as sinusoidal variations in three-phase voltage and current. We remove essentially all of these artifacts with a novel zero-phase quadratic filter.[1] This filter uses a moving window of $2w + 1$ points of raw data, $e_i$, with the same number of data points, $w$, on either side of a central point. We fit the data to a quadratic equation, $F(t_i) = a_1 T_i^2 + a_2 T_i + a_3$, with $T_i = t_i - t_c$, and $t_c$ the time at the central point of the moving window. We obtain the best fit to the data by minimizing the function, $\Psi = \Sigma_i [F(t) - e_i]^2$. The sum is over the $2w + 1$ points in the moving window. The minimum in $\Psi$ is found from the condition $\partial\Psi/\partial a_k = 0$, which yields three linear equations in three unknowns. The window-averaged signal is the fitted value at the central point, $F(t_c = t_i) = a_3$. The sums over odd powers of $T_i$ are zero; symmetric sums over even powers of $T_i$ (over $i$ from $-w$ to $w$) can be converted to sums from 1 to $w$, giving a window-averaged solution for the artifact signal,

$$F(t = t_c) = [3(3w^2 + 3w - 1)\sum_{i=-w}^{w} e_{i+c} - 15\sum_{i=-w}^{w} i^2 e_{i+c}] / (4w^2 + 4w - 3)(2w + 1). \tag{A.1}$$

The sums in this last equation are over $i$ from $-w$ to $w$, with sums over even powers of $i$ explicitly evaluated with standard formulas for $\Sigma_i i^2$ and $\Sigma_i i^4$ (ref. 2). The effort to evaluate Eq. (A.1) can be reduced further by computing the sums initially with $c = w + 1$, and then using recursions thereafter for $c > w + 1$ (ref. 1). Application of this filter to the $N$-point set of raw data, $e_i$, yields $N - 2w$ points of artifact data, $f_i = F(t_c = t_i)$. The residue, $x_i = e_i - f_i$, has essentially no artifact content. Subsequent analysis uses only the artifact-filtered machine data, $x_i$.

### A.2 CONVENTIONAL STATISTICAL MEASURES

Analysis of time serial data begins with the collection of a process-indicative scalar signal, $x$, from a dynamical system whose dimensionality, structure, parameters, and regime are usually unknown. This signal is sampled at equal time intervals, $\tau$, starting at the initial time, $t_0$, and yields a sequence of $N$ points, $x_i = x(t_0 + i\tau)$. Several linear measures are useful for characterizing the gross features of this data. The first is the mean, $\underline{x}$, or average over the $N$ data points:

$$\underline{x} = \sum_{i=1}^{N} x_i / N. \tag{A.2}$$

The second is the sample standard deviation ($\sigma$), which follows from Eq. (A.2):

$$\sigma^2 = \sum_{i=1}^{N} (x_i - \underline{x})^2 / (N-1). \tag{A.3}$$

Equation (A.3) is the second moment about the mean, implying that higher moments are available. Thus, a third linear measure is the third moment about the mean, called skewness, $s$:

$$s = \sum_{i=1}^{N} (x_i - \underline{x})^3 / N\sigma^3. \tag{A.4}$$

A fourth linear measure is the fourth moment about the mean, called kurtosis, $k$:

$$k = \sum_{i=1}^{N} (x_i - \underline{x})^4 / N\sigma^4 - 3. \tag{A.5}$$

Typical process data have significant values for skewness and kurtosis, but Gaussian random processes have values that are not significantly different from zero.[3] A large positive (negative) value of skewness corresponds to a longer, fatter tail of the data distribution about the mean to the right (left). Kurtosis measures the amount of flattening (negative $k$) or excess peakedness (positive $k$) about the mean. Another measure applies to both linear and nonlinear systems, and involves counting the number of times, $n_c$, that the signal crosses the mean value. More specifically, one-half of a wave period is delimited by two successive mean crossings. For $n_c$ >>1, the average number of time steps per wave cycle ($m$) as:

$$m = N/[(n_c - 1)/2] = 2N/(n_c - 1) \approx 2N/n_c\ . \tag{A.6}$$

This last measure indicates the average periodicity in the signal, or the inverse of the average frequency. Analysis of typical data shows that these measures provide little, if any, discrimination for detection of condition change. We include these measures for the sake of completeness and to show that linear measures are inadequate for prognostication.

## A.3 TRADITIONAL NONLINEAR MEASURES

Nonlinear analysis uses the same sequence of time serial data, $x_i$, to reconstruct the process dynamics. In particular, phase-space (PS) reconstruction[4] uses $d$-dimensional time-delay vectors, $y(i) = [x_i, x_{i+\lambda}, \ldots, x_{i+(d-1)\lambda}]$, for a system with d active variables and time lag, $\lambda$. The choice of lag and embedding dimension, $d$, determines how well the PS reconstruction unfolds the underlying dynamics from a finite amount of noisy data. Takens[5] found that, for a $d$-dimensional system, $2d + 1$ dimensions generally results in a smooth, non-intersecting reconstruction. Sauer et al.[6] showed that, using ideal data (i.e., no noise and infinite precision), the first integer greater than the correlation dimension is often sufficient to reconstruct the system dynamics; this result has been confirmed by computing the embedding dimension via the false nearest-neighbors

method.[7,8,9] However, too high an embedding dimension can result in over-fitting for real data with finite length and noise. We further note that different observables of a system contain unequal amounts of dynamical information,[10] implying that PS reconstruction could be easier from one variable, but more difficult or even next to impossible from another. Our analysis seeks to balance these caveats within the constraints of finite-length noisy data.

Various nonlinear measures have been defined to characterize process dynamics using the PS reconstruction.[11,12] We choose three of these nonlinear measures, against which we compare the dissimilarity indicators. In particular, we use: the first minimum in the mutual information function as a measure of de-correlation time, the correlation dimension as a measure of dynamic complexity, and the Kolmogorov entropy as a measure of predictability. For the reader's convenience, we briefly describe these three measures next.

The mutual information function (MIF) is a nonlinear version of the (linear) autocorrelation and cross-correlation functions and was originally developed by Shannon and Weaver[13] with subsequent application to time series analysis by Fraser and Swinney.[14] The MIF measures the average information (in bits) that can be inferred from one measurement about a second measurement and is a function of the time delay between the measurements. Univariate MIF measures predictability within the same data stream at different times. Bivariate MIF measures predictability of one data channel, based on measurements in a second signal at different times. For the present analysis, we use the first minimum in the univariate MIF, $M_1$, to indicate the average time lag that makes $x_i$ independent of $x_j$. The MIF, $I(q,r)$, and system entropy, $H$, are defined by

$$I(q,r) = I(r,q) = H(q) + H(r) - H(r,q)\,, \tag{A.7}$$

$$H(q) = -\sum_i P(q_i)\log[P(q_i)]\,, \tag{A.8}$$

$$H(q,r) = -\sum_{i,j} P(q_i,r_j)\log[P(q_i,r_j)]\,. \tag{A.9}$$

For a window of $N$ points, we denote the $Q$ set of data measurements by $q_1, q_2, \ldots, q_N$, with associated occurrence probabilities $P(q_1), P(q_2), \ldots, P(q_N)$. $R$ is a second set of measurements, $r_1, r_2, \ldots, r_N$, with a time delay relative to the $q_i$ values, with occurrence probabilities $P(r_1)$, $P(r_2), \ldots, P(r_N)$. The function $P(q_i, r_j)$ denotes the joint probability of both states occurring simultaneously. $H$ and $I$ are expressed in units of bits if the logarithm is taken in base two.

The maximum-likelihood correlation dimension, $D$, is:[15,16]

$$D = \left\{(-1/M)\sum_{ij} \ln[(\delta_{ij}/\delta_0 - \delta_n/\delta_0)/(1-\delta_n/\delta_0)]\right\}^{-1}\,, \tag{A.10}$$

where $M$ is the number of randomly sampled point pairs; $\delta_{ij}$ is the maximum-norm distance between the (randomly chosen) $i-j$ point pairs, as defined in Eq. (A.12) below. The distance (scale length) $\delta_n$ is associated with noise as measured from the time serial data. Note that the distances are normalized with respect to a nominal scale length $\delta_0$, which is chosen as a balance between sensitivity to local dynamics (typically at $\delta_0 \lesssim 5a$) and avoidance of excessive noise (typically at $\delta_0 \geq a$). Here, the symbol a denotes the absolute average deviation as a robust indicator of variability[16] in the data,

$$a = (1/N)\sum_{i=1}^{N} |x_i - \underline{x}| \tag{A.11}$$

where $\underline{x}$ is the mean of $x_i$ over the window of $N$ points. The distances $\delta_{ij}$ are defined by

$$\delta_{ij} = \max_{0 \leq k \leq m-1} |x_{i+k} - x_{j+k}|, \tag{A.12}$$

where $m$ is the average number of points per cycle, as determined by Eq. (A.6).

The Kolmogorov entropy, $K$, measures the rate of information loss per unit time, or (equivalently) the degree of predictability. Positive, finite entropy is generally considered a clear demonstration that the time series and its underlying dynamics are chaotic. A very large entropy indicates a stochastic (nondeterministic) and therefore totally unpredictable phenomenon. The K-entropy is estimated from the average divergence time for pairs of initially close orbits. More precisely, the entropy is obtained from the average time for two points on an attractor to go from an initial separation $\delta \leq \delta_0$ to a separation of more than that distance ($\delta > \delta_0$). We use the maximum-likelihood K-entropy of Schouten et al.,[17]

$$K = -f_s \log(1 - 1/\underline{b}), \tag{A.13}$$

$$\underline{b} = (1/M)\sum_{i=1}^{M} b_i, \tag{A.14}$$

with $b_i$ as the number of time steps for two points, initially within $\delta \leq \delta_0$, to diverge to $\delta > \delta_0$. The symbol $f_s$ denotes the data-sampling rate.

There are several problems associated with the use of these measures for detection of dynamical change. The most serious is that these nonlinear measures are expressed as a sum or integral over (a region of) the PS, thus averaging out all dynamical details into a single number. Two (very) different dynamical regimes may lead to very close, or even equal measures. The situation is even murkier for noisy dynamics, in which case reliable determination of the nonlinear measures is next to impossible. The second difficulty arises from the definitions of K-entropy and correlation dimension in the limit of zero scale length. However, all real data have noise and even noiseless model data is limited by the finite precision of computer arithmetic. Thus, we choose a finite scale length that is somewhat larger than the noise ($\delta_0 = 2a$), at which to report the

values of $K$ and $D$, corresponding to finite-scale dynamic structure. Consequently, the calculated values of $K$ and $D$ have smaller values than expected for the zero-scale-length limit ($\delta_0 \rightarrow 0$) and cannot capture dynamical complexity at length scales smaller than $\delta_0$. A third difficulty arises from the definition of these nonlinear measures as functionals of the distribution functions. Some of these functionals do not satisfy all the mathematical properties of a distance. In particular, for some of them, symmetry and the triangle inequality may be violated.[18] Therefore, these measures cannot define a metric in the mathematical sense. They may indicate change, although only in a sense that has to be made precise for each situation.

## A.4 PHASE-SPACE DISSIMILARITY MEASURES

The traditional nonlinear measures described in the previous section characterize global features of the dynamics, and can clearly distinguish between regular and chaotic dynamics. However, they do not reveal slight dissimilarities between dynamical states. The same is true for other global indicators, such as fractal dimension, Lyapunov exponents, etc. This lack of discrimination occurs because such traditional measures are based on averaged or integrated features of the dynamics over the attractor, which provide a global picture of long-term dynamical behavior.

Greater discrimination is possible by more detailed analysis of the reconstructed dynamics. The natural (or invariant) measure on the attractor provides a more refined representation of the reconstruction, describing the visitation frequency of the system dynamics over the PS. We obtain a useful discrete representation of the invariant measure from time serial data as follows. We first represent each signal value, $x_i$, as a symbolized form, $s_i$, that is, one of $S$ different integers, $0,1,\ldots,S-1$,

$$0 \le s_i = INT[S(x_i - x_{\min})/(x_{\max} - x_{\min})] \le S-1. \tag{A.15}$$

Here, the function (INT) converts a decimal number to the closest lower integer, and $x_{min}$ and $x_{max}$ denote the minimum and maximum values of $x_i$, respectively, over the base case (reference data). We previously used[19,20,21] the minimum and maximum values over both the base case and test case (data to be tested for departure from the base case). However, in real- or near-real-time analyses, only base case extrema are actually known. We require that $s_i(x_i = x_{max}) = S-1$ in order to maintain exactly $S$ distinct symbols. Consequently, Eq. (A.15) creates symbols that are uniformly distributed between the minimum and maximum in signal amplitude (uniform symbols).

An alternative is equiprobable symbols. These symbols are formed by ordering the base case time-serial data from the smallest to largest value. The first $N/S$ of these ordered data values correspond to the first symbol (0). Ordered data values $(N/S)+1$ through $2N/S$ correspond to the second symbol (1), and so on up to the last symbol, $S$-1. Consequently, equiprobable symbols have non-uniform partitions in the signal amplitude so that each symbol has the same occurrence frequency $(N/S)$ of $x_i$ values. Much structure is inherent in uniform symbols before beginning the PS reconstruction, but no PS structure arises from equiprobable symbols. Thus, a key advantage of equiprobable symbols is that dynamical structure arises only from the phase-space

reconstruction, as described below. Large negative and large positive values of $x_i$ have little affect on equiprobable symbolization, but dramatically change the partitions for uniform symbols. Moreover, information theoretic measures of the PS-DF (e.g., mutual information function) are a smooth function of the reconstruction parameters for equiprobable symbols, but are noisy functions of these same parameters for uniform symbols. We find that equiprobable symbols provide better discrimination of condition change than uniform symbols.

The phase-space is partitioned into $S^d$ hypercubes or bins by the symbolization process. We then count the number of PS points occurring in each bin to obtain the distribution function (DF) as a discretized density on the attractor. We denote the population of the ith bin of the distribution function, $Q_i$, for the base case, and $R_i$ for a test case, respectively. For infinitely precise data, this representation has been used in Grebogi et al.[22] The choice of parameters ($S$, $N$, and $d$) depends not only on the system, but also on the specific data under consideration. In the preliminary phase of the analysis, we systematically varied each parameter with the others fixed, to obtain optimum sensitivity of the measures to changes in system dynamics for each class of data. After achieving optimal sensitivity, the values of the parameters were kept fixed.

Initial analysis used an embedding window, $M_I = (d - 1)\lambda$, based on the first minimum in the mutual information function, $M_I$ (ref. 14). This choice of time delay provides maximal information for the reconstruction of the phase space dynamics. Then, we set $\lambda = \mathrm{INT}[0.5 + M_I/(d - 1)]$ to obtain an integer value for the lag when $M_I$ is not evenly divisible by $d - 1$. The reconstruction requires that $\lambda \geq 1$, thus constraining the largest value of dimensionality to $d \leq 2M_I + 1$ from the above formula. Subsequently, we have found that this choice of time-delay lag is not the best for failure forewarning. Rather, we vary $\lambda$ as one of several parameters that determine the goodness of the PSDM in providing forewarning of failure.

After reconstruction (unfolding) of the dynamics, the test case is compared to the base case. Diks et al.[23] measured differences between delay vector distributions by the square of the distance between two DFs. Schreiber[24,25] measured dissimilarity via the Euclidean distance between points of the attractor. These measures of dissimilarity only account for the geometrical shape and location of the attractor. Manuca and Savit[26,27] described dissimilarity via ratios of the correlation integral over the DF. This is essentially the correlation dimension, as discussed above. Moreover, these papers discuss dissimilarity measures from the perspective of non-stationarity, while our focus is on quantification of condition change. In particular, we measure the difference between $Q_i$ with $R_i$ by the $\chi^2$ statistics and $L_1$ distance,

$$\chi^2 = \sum_i (Q_i - R_i)^2 / (Q_i + R_i), \tag{A.16}$$

$$L = \sum_i |Q_i - R_i|, \tag{A.17}$$

where the summations in both equations run over all of the populated PS cells. The choice of these measures is based on the following considerations. The $\chi^2$ statistic is one of the most powerful, robust, and widely used statistical tests to measure discrepancies between observed and expected frequencies. The $\chi^2$ statistic is obviously symmetric, but does not always

satisfy the triangle inequality, so it does not define a distance in the mathematical sense. The $L_1$ distance is the natural metric for distribution functions since it is directly related to the total invariant measure on the attractor and does define a bona fide distance. Therefore, these measures account for changes in the geometry, shape, and visitation frequency of the attractor and can be viewed as somewhat complementary. Obviously, calculation of these measures in a consistent fashion, requires that the base case and test case contain the same number of points, identically sampled; otherwise the distribution functions have to be properly rescaled.

We extended the previous analysis in a manner that is naturally compatible with the underlying dynamics. By connecting successive PS points as prescribed by the dynamics, $y(i) \rightarrow y(i+1)$, we obtain a discrete representation of the process flow.[28] Thus one can form a $2d$-dimensional vector, $Y(i) = [y(i), y(i+1)]$, by adjoining two successive vectors from the $d$-dimensional reconstructed PS, lives in a $2d$-dimensional space, that we call the connected phase space (CPS). As before, $Q$ and $R$ denote the CPS DFs for the base case and test case, respectively. We define the measures of dissimilarity between these two CPS DFs, as before, via the $L_1$-distance and $\chi^2$ statistic,[19,21,29,30]

$$x_c^2 = \sum_{ij} \left(Q_{ij} - R_{ij}\right)^2 / \left(Q_{ij} + R_{ij}\right) \tag{A.18}$$

$$L_c = \sum_{ij} | Q_{ij} - R_{ij} | . \tag{A.19}$$

The subscript $c$ indicates the connected distribution function measure. We note that the value $\lambda = 1$ results in $d - 1$ components of $y(i+1)$ being redundant with those of $y(i)$, but we allow this redundancy to accommodate other data such as discrete points from two-dimensional maps. The CPS measures have a higher discriminating power than their non-connected counterparts. Indeed, we can prove that the measures defined in Eqs. (A.16)–(A.19) satisfy the four inequalities[29] $\chi^2 \leq L$, $\chi_c^2 \leq L_c$, $L \leq L_c$, and $\chi^2 \leq \chi_c^2$. Alternative forms are: $\chi^2 \leq L \leq L_c$ and $\chi^2 \leq \chi_c^2 \leq L_c$.

The $\chi^2$ statistic requires statistical independence between various samples. However, the PS points depend on one another due to reconstruction from time delay vectors with dynamical structure.[23] The resulting statistical bias is avoidable by averaging contributions to Eqs. (A.16)–(A.19) over values of y(j) or Y(j) which satisfy $|i - j| < \Lambda$ (ref. 23), where $\Lambda$ is some largest typical correlation time lag. We tested the bias in typical data by sampling every $\Lambda$-th connected phase space point for $4 \leq \Lambda \leq 23$, resulting in $\Lambda$ different samples for the base case ($Q_i$) and for each cutset ($R_i$). We then averaged the sampled $\chi^2$ values over the $\Lambda^2$ different combinations of distribution functions for the base case and test case cutsets. As expected, a decrease proportional to $1/\Lambda$ occurs in the sampled $\chi^2$ values, because the number of data points contributing to $\chi^2$ decreases in the same proportion. The trend over time in sampled $\chi^2$ values is the same as in $\chi^2$ values without sampling, showing that no bias is present. Thus, we use unsampled $\chi^2$ values for the remainder of this work as a relative measure, rather than as an unbiased statistic for accepting or rejecting a null statistical hypothesis.[19]

Use of the dissimilarity measures on finite length, noisy data requires a consistent statistical implementation and interpretation. We use the first B non-overlapping cutsets as base cases. The choice of the number of basecase datasets, $B$, should strike a judicious balance between a reasonably short base case period to capture quasi-stationary, "normal" dynamics and a sufficiently long period for statistical significance. We have chosen $B = 10$ for noisy machine data to provide a sufficient statistical sample.

The disparate range and variability of various nonlinear measures are difficult to interpret (especially for noisy data), so we need a consistent means of comparison. Thus, we renormalize the nonlinear measures.[19,21] For each nonlinear measure, $V = \{D, K, M_l, L, L_c, \chi^2, \text{and } \chi_c^2\}$, we define $V_i$ as the value of the nonlinear measure for the $i$th cutset. As before, $V$ is the mean value of that nonlinear measure over the non-outlier base cases, with a corresponding sample standard deviation $\sigma$, as described above. No averaging is needed for $D$, $K$, and $M_l$ since the calculation of these measures involves only one cutset at the time. The renormalized form is then $U(V) = |V_i - V|/\sigma$, which measures the number of standard deviations that the test case deviates from the base case mean. Several successive occurrences above threshold provide a clear indication of condition change. Alternatively, a systematic rise in the PSDM will indicate a clear departure from the base case dynamics, and provides forewarning of failure.

## REFERENCES


1. L. M. Hively, N. E. Clapp, C. S. Daw, and W. F. Lawkins, "Nonlinear Analysis of EEG for Epileptic Events," ORNL/TM-12961 (Oak Ridge National Laboratory, Oak Ridge, TN, 1995).

2. I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Series, and Products, Academic Press publ., New York (4th ed., 1965).

3. M. Abramowitz and I. A. Stegun (ed.), Handbook of Mathematical Functions, U.S. Government Printing Office, Washington, D.C. (1965), page 930.

4. J. P. Eckmann and D. Ruelle, "Ergodic Theory of Chaos and Strange Attractors," Rev. Mod. Phys. 57 (1985) 617–55.

5. F. Takens, "Dynamical Systems and Turbulence," lecture notes in *Mathematics*, **898** (1981) 366 (Springer publ., New York).

6. T. Sauer, J. Yorke, and M. Casdagli, "Embedology," *J. Stat. Phys.*, 65 (1991) 579–616.

7. H. D. I. Abarbanel, R. Brown, J.J. Sidorowich, and L. S. Tsimring, "The Analysis of Observed Chaotic Data in Physical Systems," *Rev. Mod. Phys.*, 65 (1993) 1331–1392.

8. H. D. I. Abarbanel and M. B. Kennel, "Local False Nearest Neighbors and Dynamical Dimensions from Observed Chaotic Data," *Phys. Rev. E*, 47 (1993) 3057–3068.

9. L. Cao, "Practical Method for Determining the Minimum Embedding Dimension for a Scalar Time Series," *Physica D*, **110** (1997) 4–59.

10. C. Letellier, J. Maquet, L. Le Sceller, G. Gouesbet, and L. A. Aguirre, "On the Non-equivalence of Observables in Phase-space Reconstructions from Recorded Time Series," *J. Phys. A.*, **31** (1998) 7913–7927.

11. H. Kantz and T. Schreiber, Nonlinear Time Series Analysis, (Cambridge University Press) 1997.

12. A. Rezek and S. J. Roberts, "Stochastic Complexity Measures for Physiological Signal Analysis," *IEEE Trans. Biomedical Eng.* **45** (1998) 1186–1191.

13. C. E. Shannon and W. Weaver, The Mathematical Theory of Communication (Univ. of Illinois Press, Urbana) 1949.

14. A. M. Fraser and H. L. Swinney, "Independent Coordinates for Strange Attractors from Mutual Information." *Phys. Rev. A*, **33** (1986) 1134–40.

15. F. Takens, "On the Numerical Determination of the Dimension of an Attractor," lecture notes in *Mathematics*, **1125** (1984) 99–106 (Springer publ., Berlin).

16. J. C. Schouten, F. Takens, and C. M. van den Bleek, "Estimation of the Dimension of a Noisy Attractor," *Phys. Rev. E*, **50** (1994a) 1851–1861.

17. J. C. Schouten, F. Takens, and C. M. van den Bleek, "Maximum Likelihood Estimation of the Entropy of an Attractor," *Phys. Rev. E*, **49** (1994b) 126–129.

18. R. Quian Quiroga, J. Arnhold, K. Lehnertz, and P. Grassberger, "Kuhlback-Leibler and Renormalized Entropies: Applications to Electroencephalograms of Eepilepsy Patients," *Phys. Rev. E.* **62** (2000) 8380–8386.

19. L. M. Hively, P. C. Gailey, and V. A. Protopopescu, "Detecting Dynamical Change in Nonlinear Series Time Series," *Phys. Lett. A*, **258** (1999) 103–114.

20. P. C. Gailey, L. M. Hively, and V. A. Protopopescu, "Robust Detection of Dynamical Change in Scalp EEG," *Proc. Fifth Experimental Chaos Conference*, June 28–July 1, 1999, Orlando, Florida.

21. L. M. Hively, P. C. Gailey, and V. A. Protopopescu, "Sensitive Measures of Condition Change in EEG Data," *Proc. International Workshop: ''Chaos in Brain?,''* (Bonn, Germany, March 10–12, 1999), ed. K. Lehnertz et al. (World Scientific, Singapore, 2000), p. 333–336.

22. C. Grebogi, E. Ott, and J. A. Yorke, "Unstable Periodic Orbits and the Dimensions of Multifractal Chaotic Attractors," *Phys. Rev. A*, **37** (1988) 1711–1724.

23. C. Diks, W. van Zwet, F. Takens, and J. DeGoede, "Detecting Differences Between Delay Vector Distributions," *Phys. Rev. E*, **53** (1996) 2169–2176.

24. A. T. Schreiber, "Detecting and Analyzing Nonstationarity in a Time Series using Nonlinear Cross Predictions," *Phys. Rev. Lett.*, **78** (1997) 843–846.

25. A. T. Schreiber, "Interdisciplinary Application of Nonlinear Time Series Methods," *Phys. Rep.*, **308** (1999) 1–64.

26. R. Manuca and R. S. Savit, "Stationarity and Nonstationarity in Time Series Analysis," *Physica D*, **99** (1996) 134–161.

27. R. Manuca and R. S. Savit, "Nonstationarity in Epileptic EEG and implications for Neural Dynamics," *Math. Biosci.*, **147** (1998) 1–22.

28. H. D. I. Abarbanel, Analysis of Observed Chaotic Data (Springer pub., New York) 1996.

29. L. M. Hively, V. A. Protopopescu, and P. C. Gailey, "Timely Detection of Dynamical Change in Scalp EEG Signals," *Chaos* **10** (2000a) 864–875.

30. V. A. Protopopescu, L. M. Hively, and P. C. Gailey "Epileptic Event Forewarning from Scalp EEG," *J. Clin. Neurophysiol.* **18** (May 2001) 223–245.

PENNSTATE ARL

# Test Plan for Mechanical Diagnostics Test Bed in Support of NERI2000-109

## To

Dr. Lee M. Hively

Oak Ridge National Laboratory
P.O. Box 2009
Oak Ridge, TN 37831-8006

February 27, 2002

## By

Dr. Karl M. Reichard
Ms. Terri Merdes

Pennsylvania State University
Applied Research Laboratory
P.O. Box 30
State College, PA 16804
(814) 863-7681 (Voice)
(814) 863-0673 (FAX)
kmr5@psu.edu

## Introduction

This appendix describes testing that was performed on the PSU (Pennsylvania State University) Mechanical Diagnostics Test Bed (MDTB) in support of work at Oak Ridge National Laboratory (ORNL) under the U.S. Department of Energy NERI2000-109 (Nuclear Energy Research Initiative) project. This test plan describes the MDTB, the equipment that will be tested, instrumentation and data acquisition equipment, and the conditions and procedure under which the equipment will be tested.

## Test Bed

Figure B.1 shows the MDTB, which was built as an experimental research station for the study of fault evolution in mechanical gearbox power transmission components. The gearbox is driven at a set input speed using a 30 HP, 1750 RPM AC (drive) motor. A mechanical load (torque) is applied to the gearbox by a 75 HP, 1750 RPM AC (absorption) motor. The maximum speed and torque are 3500 RPM and 225 ft-lbs respectively. Speed variation is accomplished by varying the frequency to the motor with a digital vector drive unit. The variation of the torque is accomplished by a similar vector unit capable of controlling the current output of the absorption motor. The system speed and torque set points are produced by analog input signals (0-10 VDC) supplied by the data acquisition controller computer. The MDTB is capable of parallel or right angle gear motor mounts.

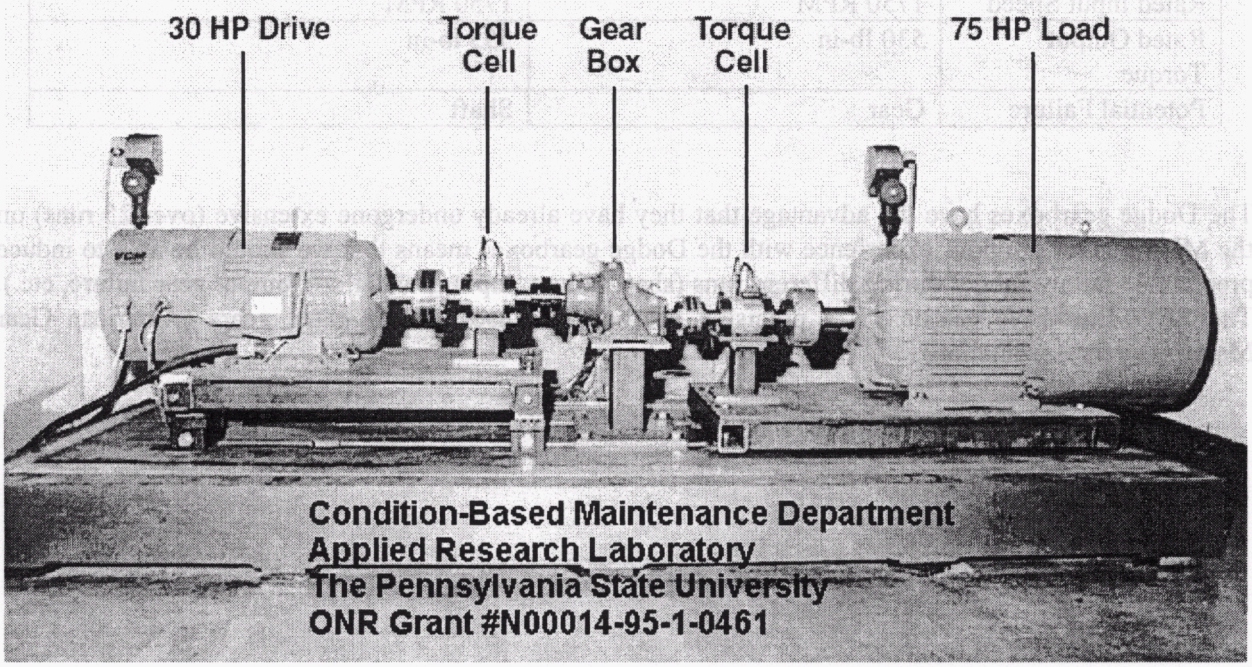



**Figure B.1: Mechanical Diagnostics Test Bed**

The MDTB has the capability of testing single and double-reduction industrial gearboxes with gear ratios from about 1.2:1 to 6:1 and with ratings that can range from 5 to 20 HP. Duty cycle profiles can be prescribed for varying speed and load. Drive line speeds for tests to date have been fixed at 1750 RPM with variable load profiles that step up to maximum values of 2 to 5 times the rated torque of the test

gearbox. The motors and gearbox are hard-mounted to minimize vibration transmission, and are precision aligned using laser technology. The shafts can be connected with either flexible or rigid couplers. Torque limiting clutches are used on both sides of the gearbox excessive torque that could occur with gear jamming or bearing seizure. Also, torque cells are used on both sides of the gearbox to directly monitor the loads transmitted and efficiency. The vector drives control the drive and load motors, providing output signals, which are sampled and stored. Output data include: input power to the drives, root-mean-square (RMS) currents, winding temperatures, motor speed, and generator torque. These signals allow automation and shutdown of motors directly through the controller PC.

## Equipment to be Tested

The MDTB was designed to perform run-to-failure testing on gearboxes. To date, only single-reduction gearboxes have been tested on the MDTB. The intent under this program is to use the MDTB to collect data on single-reduction gearboxes. Table B.1 describes the characteristics of two candidate gearboxes. The first two sets of data from the MDTB will be collected on Dodge gearboxes. Future sets may be collected using the SEW[1] Eurodrive gearboxes. Table B.2 and Table B.3 show the bearing and gear mesh frequencies for the Dodge gearbox at rated input speed.

Table B.1: Candidate Gearboxes

| Brand | Dodge (R86001) | Dodge (R86005) |
|---|---|---|
| Model Number | APG Size 3 | APG Size 3 |
| Description | Single Reduction Helical | Single Reduction Helical |
| Ratio | 1.5 | 3.38 |
| Rated Input Speed | 1750 RPM | 1750 RPM |
| Rated Output Torque | 530 lb-in | 555 lb-in |
| Potential Failure | Gear | Shaft |

The Dodge gearboxes have the advantage that they have already undergone extensive (over 25 runs) on the MDTB. Our previous experience with the Dodge gearboxes means that we should be able to induce predictable failure modes during different runs (shaft failure, input gear failure, output gear failure, etc.). The SEW Eurodrive gearbox is a precision gearbox that uses AGMA-rated gears (American Gear Manufacturers Association).

[1] SEW (Süddeutsche-Elektromotoren-Werke) was the name of the original German company, which was founded by Christian Pähr in 1935.

**Table B.2: Dodge Gearbox Bearing Frequencies (at rated input speed)**

| Bearing Input and Output Frequencies – 1.5 & 3.3 Gearbox Ratio | | | | |
|---|---|---|---|---|
| Description | **Ball Bearing (outer)** | **Ball Bearing (inner)** | **Tapered Roller Bearing (inner)** | **Taper Roller Bearing (outer)** |
| **Part Number** | **6307** | **6309** | **#15520/15578** | **#2520/2581** |
| **Input Freq (Hz)** | 29.1667 | 29.1667 | 16.0715 | 16.0715 |
| **(OUTFREQ) FTF (Hz)** | 10.76 | 11.07 | 8.03 | 7.85 |
| **(OUTFREQ) BPFO (Hz)** | 86.06 | 88.57 | 128.56 | 126.08 |
| **(OUTFREQ) BPFI (Hz)** | 147.27 | 144.77 | 175.97 | 178.46 |
| **(OUTFREQ) BSF (Hz)** | 51.77 | 57.03 | 59.64 | 52.82 |

**Table B.3: Dodge gearbox Gear Mesh Frequency (at rated input speed)**

| **Gear Frequency** | |
|---|---|
| Gearbox Size 3 – Ratio 1.5 | 875.5 Hz |
| Gearbox Size 3 – Ratio 3.3 | 613.0 Hz |

## Instrumentation and Data Acquisition Equipment

Data will be collected using a National Instruments PXI measurement system. Figure B.2 shows a picture of an example NI system. This data acquisition system is composed of a backplane, a processor and control module, and separate data acquisition modules. The MDTB testing will use NI4472 dynamic signal acquisition and analysis modules (Figure B.3). Module operating characteristics are given in Table B.4. Each module has 8 analog inputs with simultaneously-sampled, 24-bit, sigma-delta A/D converters. The maximum sample rate for each channel is 102.4 kHz. The internal digital anti-aliasing filters in the sigma-delta A/D's are designed to have the following dynamic characteristics (per National Instruments specification sheet for the data acquisition module):

Alias-free bandwidth (passband): DC to 0.4535 f;
Stop band: 0.5465 f;
Alias rejection: 110 dB.

These specifications are consistent with comparable sigma-delta A/D systems. We intend to acquire data at 51.2 kHz sample rate, which should provide an alias-free bandwidth of DC to 23 kHz. Note that the mounted resonance frequency of the accelerometers that will be used for vibration measurements is above 70 kHz, well outside the bandwidth of the proposed measurements.

The following data will be collected:

3-phase input motor voltages
3-phase input motor currents
3-axis acceleration measurement on gearbox housing
input and output torque

All channels will be sampled at approximately 50 kHz in 10 second snapshots. The data will be saved to a computer hard disk drive. Operating specifications for the tri-axial accelerometer are given in Table 5. Additional sensor measurements may be added on open data acquisition channels to provide consistency with earlier gearbox tests. The full sensor list will be provided in the post-run test description along with a drawing showing sensor placement.

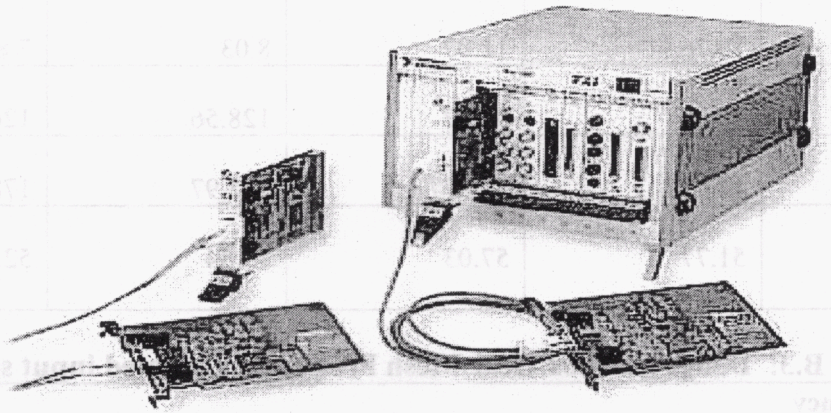


**Figure B.2: National Instruments PXI Data Acquisition System**

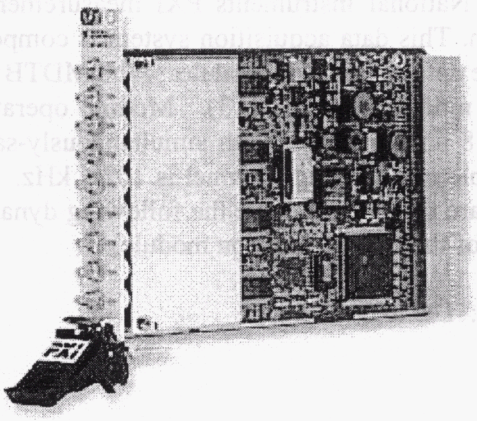


**Figure B.3: National Instruments Dynamic Data Acquisition Module**

**Table B.4: DAQ Module Specifications**

| | |
|---|---|
| Channels per module | 8 |
| A/D resolution | 24 bits |
| Dynamic range | 120 dB |
| Measurement bandwidth | DC-45 kHz |
| Coupling | AC/DC |

**Table B.5: Accelerometer Specifications**

| | |
|---|---|
| Sensor Name | Accelerometer |
| Sensor Make | PCB Piezotronics, Inc |
| | Quartz Shear ICP |
| Sensor Model # | 356B08 |
| Sensor Serial # | 8052 |
| Sensor Type | Shear Piezoelectric |
| Sensor Volt Sensitivity | 100 mV/g |
| Measurement Range | ±50 g |
| Frequency Range (±5%) | 1 to 10 000 Hz |
| Mounted Resonant Frequency | > 70 kHz |
| Broadband Resolution | 0.005 g rms (0,05 m/s2 pk) |
| | |
| | |
| Conditioner Make | PCB Piezotronics, Inc. |
| Conditioner Model # | 481A02 |
| Conditioner Gain | 2 |

## Data Format

Data will be provided in a documented binary format. MatLab m-files were also provided to read the header information and data from the files. Data were delivered to ORNL on either CDR or DVDR. The anticipated total size for a one, 10s snapshot containing 11 sensor channels is 22.5 MB (10 seconds x 51.2 k samples/s x 4 bytes/sample/channel x 11 channels/snapshot). If each channel is saved into a separate file, the size per channel per snapshot would be roughly 2.2 MB per file.

## Test Conditions and Procedure

The test procedure that will be followed is described below:
1) Disassemble gearbox and drain out oil;
2) Take images of both pinion and gear teeth;
3) Index gear if possible;
4) Place tooth identification numbers on gear;
5) Mark mating pinion/gear teeth with a large slash;
6) Assembly gearbox using mating slash and refill with manufacturer gear oil;
7) Mount the gearbox to the test stand and laser align;
8) Attach all sensors to the gearbox and set up data acquisition system;
9) Calibrate all sensors and save results, with time stamps to correlate with oil data;
10) Run Test Matrix until failure;
11) Disassemble gearbox and take images of both pinion and gear teeth;
12) Index gear if possible.

The intended test conditions are shown below in Table B.6. Data will be collected until failure of the gearbox or until the damage to the gearbox threatens to induce damage in other system components.

**Table B.6: Test Condition Matrix**

| Test Condition | Speed (RPM) | Torque | Length of Condition | Snapshot Rate | Ground Truth Information |
|---|---|---|---|---|---|
| 1 | 1750 | 100% of rated | 1 hour | 15 mins | Send data to ORNL for quality check<br>ARL verify quality of data<br>Oil sample<br>Borescope gears |
| 2 | 1750 | 300% of rated | Until failure | 30 mins** (nom.) | Borescope gears if changes are detected via algorithms<br>Oil samples every 24 hours<br>Send data to ORNL for quality check and analysis |

** Sample rates will change from 30 minutes to 1-10 minute intervals as algorithms detect changes.

The test protocol involves: (i) a linear increase in motor speed from zero to the operational speed, (ii) maintenance of this constant speed for some chosen period, (iii) a linear decrease in motor speed from (ii) to zero, (iv) sampling of the gearbox oil for particulates as a measure of gear wear. The cycle of (i)-(iv) was repeated many times under the test load until the gearbox failed, causing excessive vibration, which triggered termination of the experiment. ORNL analysis uses only the flattop data from (ii) of the test cycle.

## Appendix C: Test Plan for Torsion Experiment

PENNSTATE

1855

**Title:** Seeded Crack Fault Test of Laboratory Bladed Disk Assembly

**Personnel:** Brian Resor, Martin Trethewey, Ken Maynard

# Test Description

One failure mode in a turbo-machine begins with a crack at the base of a rotating blade, eventually causing blade loss. This Appendix describes an experiment to simulate such a failure. PSU conducted the experiment on the Torsional Vibration Test Rig during May of 2002. Figure C.1 shows the test rig. The objective was detection of dynamical changes with increasing crack size, thus simulating the change in dynamical frequencies due to crack initiation and growth.

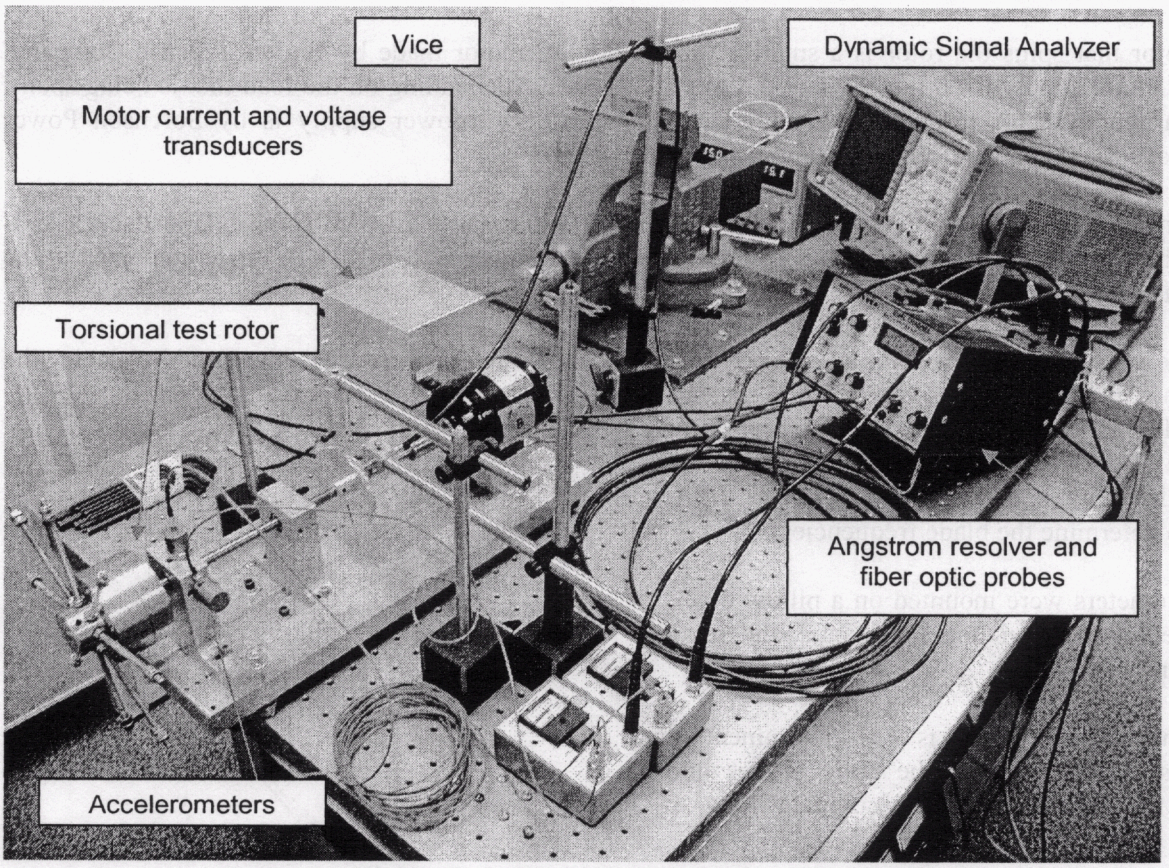



Figure C.1. Motor testing equipment setup

PSU work to date has simulated the crack growth by changing the lock-nut locations along the threaded rods. Figure C.2 shows the rotor assembly, which has eight equally spaced threaded rods to simulate the blades. A fractional horsepower DC motor rotates the blades. The present experiment is an extension of this earlier PSU work, involving a sequence of tests with a progressively deeper machined "crack" to change the rod frequencies, instead of moving the lock nuts. The seeded crack was placed in one of the threaded rod "blades" by using a wire electric discharge machining (EDM) to cut a 0.010 in wide slot. Test data at each depth of cut included motor voltage, motor current, two-axis translational vibration on one bearing pillow block, and torsional vibration of the rotor shaft. A description of the data files is shown at the end of this Appendix. Additional details can be found in the 2002 PSU Masters Thesis by Brian Resor.

The test protocol was as follows:

1. Acquire test data (items 3-5, below) for the no-cut (nominal) state of the rotor assembly.
2. Place an initial 0.010-inch cut in one rod.
3. Measure the bending natural frequency of the seeded fault rod.
4. Place the rotor assembly in the torsional vibration test stand.
5. Run the test stand and acquire the time-serial data from the system sensors.
6. Remove the rotor assembly and increment the slot depth another 0.010 inch by EDM.
7. Repeat steps 2 and 6 until a "failed" state is achieved (after six successive EDM cuts).

Deliverables for this experiment include:

1. Data for each of the seven tests of the rotor (one nominal state, plus six cuts).
2. Experimental characterization of each test state.

The motor that spins the rotor is a small 10,000 RPM DC motor made by Bodine Electric Company. Typical motor supply is on the order of 4 Volts and 2 Amps (depending on the load that is being spun). This particular motor is not manufactured anymore. The DC power supply is by Sorenson Power Supplies, a Raytheon Company (part number DCR150-12B).

The transducer that was used for measuring motor input voltage was a LEM Voltage Transducer CV 3-200. The current was measured using a LEM Current Transducer LTS 6-NP. Both are contained in the box that is pictured in Fig. C.1.

The vice is used to hold the rotor assembly, while the "blade" static frequency is measured by placing the tip of the fiber optic probes very close to the end of the blade and plucking the blade. The probes sense the size of the gap between their tip and the tip of the blade. This signal is analyzed in the DSA in order to determine the blade frequencies.

Accelerometers were mounted on a pillow block to measure vibration in both horizontal and vertical directions.

The angstrom resolver is the instrument that converts the signal from the optic probes into a voltage which is then fed to the computer.

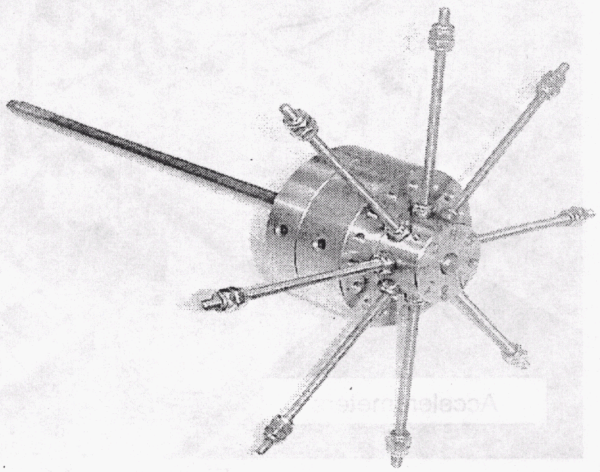


**Figure C.2.** Simulated bladed disk assembly

The data acquisition system is a Hewlett Packard VXI Mainframe with an E1433A 8-channel data acquisition board with tachometer inputs. Data is sent from the VXI Mainframe to a desktop PC using the HP E8491A firewire card.

The desktop PC used a software package called HP DAQ Express to manage the data acquisition. Using this software along with the E1433 card, it was possible to acquire 3 simultaneous channels of data at 12.8 KHz sample rate.

## Results of Crack Simulation

The seeded defect for these tests was a small "crack" at the base of one of the eight blades of the rotor, as shown in Fig. C.3. Wire EDM cutting was used to produce the smallest possible cut to simulate a crack. The diameter of the wire that was used for these cuts was 0.010 inches and the over burn is approximately 0.001 inches. This adds up to a total cut width of 0.012 inches. The cut depth can be controlled to within about 0.0005 inches.

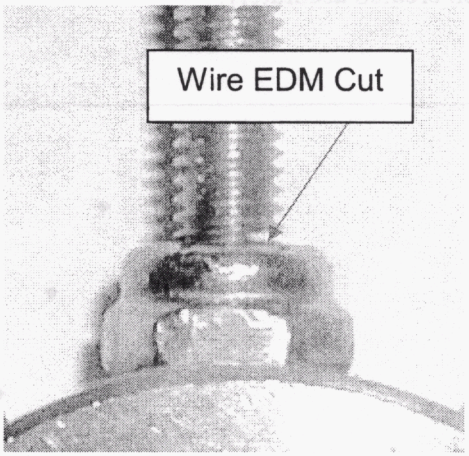



Figure C.3. Picture of blade cut

The cut location was as close to the blade root as was practical. A fixture was created for use in the Wire EDM machine to which the whole bladed assembly (excluding shaft) is mounted. Using computer controlled tooling, the cut location and depth can be carefully controlled for each cut. The depth of the first cut was measured from the point that the 0.010 inch wire came in contact with thread surfaces at the deepest point of the thread.

The original tuned frequency of all the blades was set at 205 Hz (within +/- 0.25 Hz). Cuts were made to each of the depths found in Table C.1 and the static frequencies were recorded. Static frequencies for first bending modes in both axes were recorded. The "soft" static frequency corresponds to the frequency that couples with shaft torsion and is measured by the torsional vibration measurement technique.

Table C.1: Torsional test results summary

| Cut | spec'd cut depth (in) | Soft static blade frequency (Hz) | Stiff static blade frequency (Hz) | Coupled Mode | 1st Shaft Mode | Dynamic Rogue Blade Frequency | Speed (RPM) | Coupled Mode | 1st Shaft Mode | Dynamic Rogue Blade Frequency | Speed (RPM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 205 | 205 | 213.5 | 268.5 | not visible | 2941 | 195 | 233 | not visible | 2930 |
| 1 | 0.024 | 203 | 205.25 | 213.5 | 266.5 | not visible | 2858 | 198.5 | 240.75 | 207.25 | 3032 |
| 2 | 0.032 | 201.75 | 205 | 213.5 | 265.5 | 206 | 3019 | 198.5 | 239.5 | 206 | 3014 |
| 3 | 0.037 | 201.5 | 205.25 | 214 | 264.25 | 205 | 2988 | 198.25 | 240 | 206 | 3039 |
| 4 | 0.042 | 198.5 | 204.5 | 213 | 264.5 | 203 | 3029 | 197.25 | 235.25 | 203.5 | 2986 |
| 5 | 0.047 | 195 | 203.75 | 212.5 | 264.25 | 200 | 2959 | 197 | 241 | 201.75 | 3022 |
| 6 | 0.049 | 193.75 | 203.5 | 211.5 | 261.8 | 199.1 | 2994 | 193.3 | 233.8 | 200.7 | 3004 |

Figure C.4 shows the relationship between cut depth ratio and percentage degradation in blade frequency. Cut depth ratio is defined as cut depth L to pitch diameter D of the threaded rod (D is about 0.016 inches). Note that the trend in frequency degradation very nicely follows a definite trend except for one point that corresponds to Cut 3. The rig was probably not mounted quite right for this cut and therefore the cut depth might not have been created accurately.

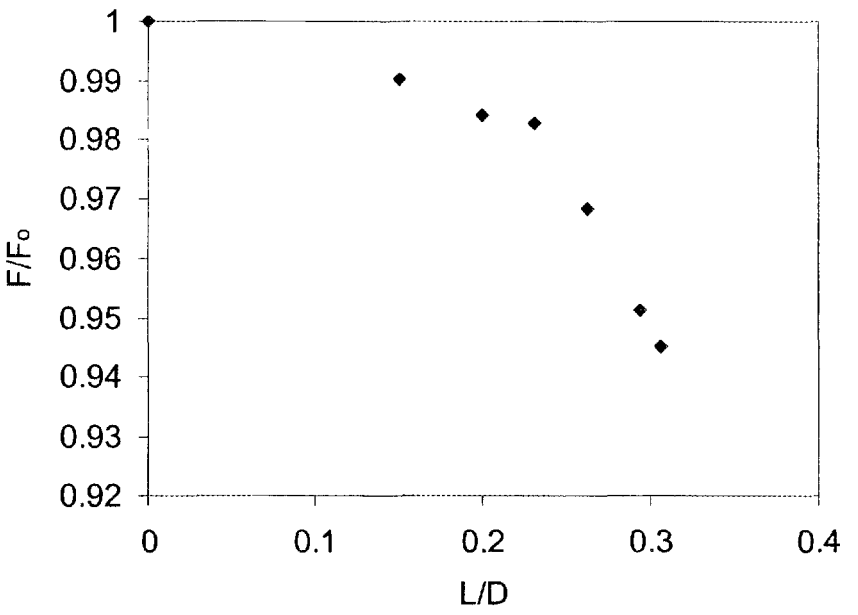



Figure C.4. Cut depth versus frequency change

The rotor was assembled in the torsional rig and it was spun at approximately 2900-3000 rpm. The running speed was maintained approximately the same for all tests so that effects of aerodynamic excitation and centripetal speed stiffening are equal throughout.

In some of the motor voltage and current and acceleration data there may be evidence of speed fluctuations. During some of the testing it was difficult to keep the rig running at a constant speed for some unknown reason. The motor speed would often cycle by about 50 RPM every 15-20 seconds. This problem was not experienced in the past.

## Torsional Measurements

Torsional data was acquired by detecting zero crossings from a striped tape encoder that is installed on the shaft of the rotor. Measured zero crossings are compared to imaginary zero crossings that would be observed with no torsional vibration in order to determine the phase shift of the carrier wave. When rotating speed is known, shaft twist is directly related to phase shift. See the 2002 PSU Masters Thesis by Brian Resor for more information on this technique.

The torsional data from this testing is the actual shaft twist time waveform in the form of degrees of twist versus time. The times that are reported are the exact sample times (which are not spaced constantly due to the nature of this technique). The shaft twist is calculated from the measured zero crossings from a 159 pulse-per-revolution striped tape. The data is corrected for errors that are present in the striped tape. A 1st order digital Butterworth high-pass filter with 50 Hz cutoff is also applied in order to remove the effects of very gradual shifts in running speed.

The torsional spectra for this experiment are the logarithm of torsional displacement versus frequency. For example, a value of –3.5 on the vertical scale corresponds to a peak vibration amplitude of $10^{-3.5}$ degrees = 0.000316 degrees.

## Torsional Results

For each cut depth, data was acquired with the rig in two different configurations. Setup 1 is the typical setup that has been studied in all the previous work on torsional vibration monitoring for turbine blade health. The location of the coupled mode (See Figure C.5) is 5-10 Hz higher than the rogue blade mode. In this configuration, as a rogue blade deteriorates it appears to separate and grow away from the large coupled mode peak. Figure C.6 shows the torsional results of this portion of the test.

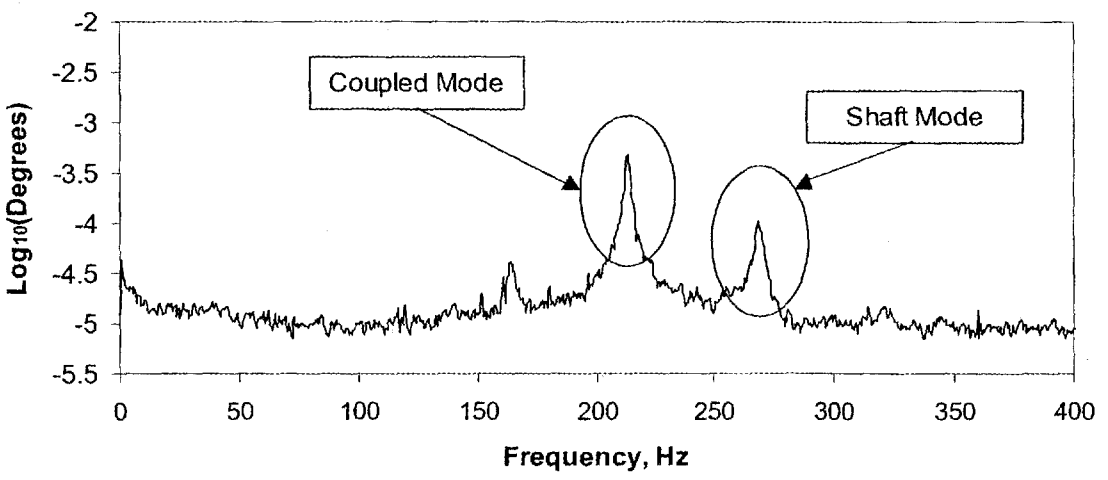



Figure C.5: Important torsional peak locations

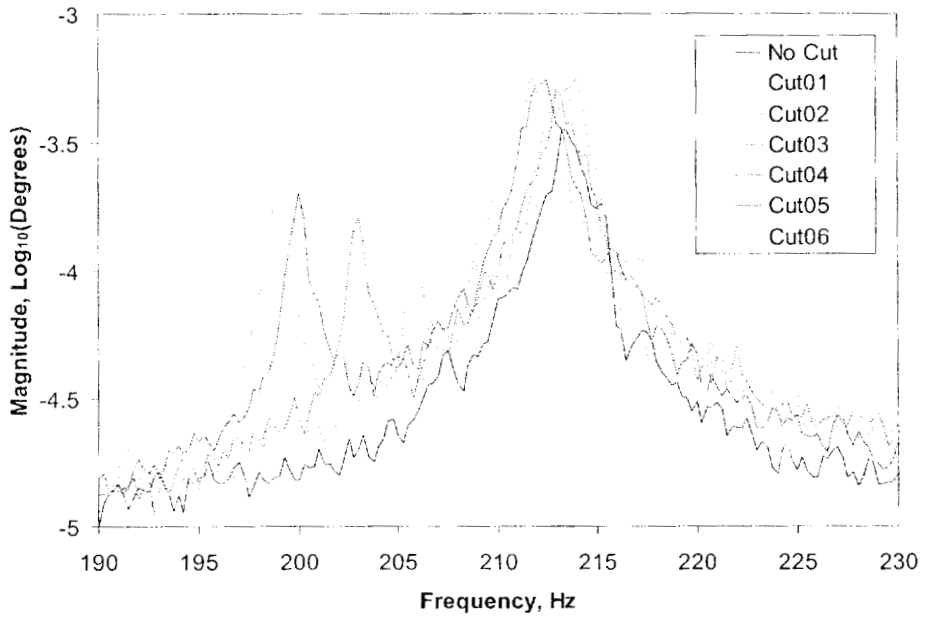



Figure C.6: Rogue blade cascade for Setup 1

Setup 2 is a different configuration in that the coupled mode is about 10 Hz lower than the rogue blade mode. This configuration is achieved by simply moving the masses on the shaft by a small amount.

Results for testing in Setup 2 are found in C.7.

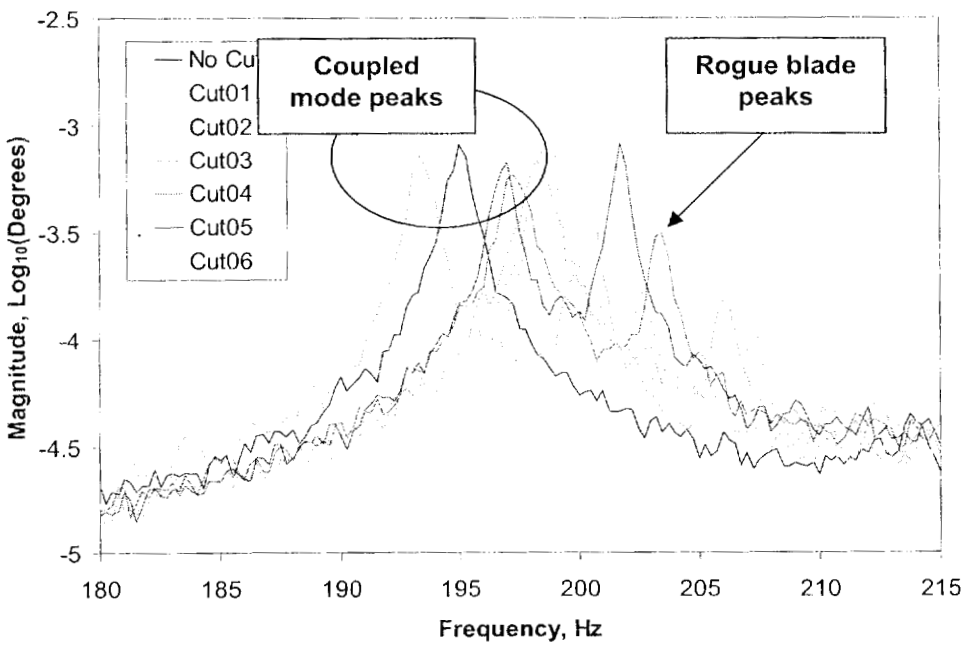



Figure C.7. Rogue blade cascade for Setup 2

The coupled-mode peak appears in slightly different locations in each case due to difficulties associated with disassembly and reassembly of the shaft components for each cut. These effects are insignificant. This experiment shows that the rogue blade mode behaves as expected, moving away from the tuned frequency die to coupling with the shaft torsion. The rogue blade vibration amplitude increases as the frequency change increases. Moreover, this test shows that a very fine cut shifts the rotor-blade frequency that is detectable via the torsional vibration measurement.

## Data File Descriptions

PSU provided four MatLab™-formatted binary data files for each wire EDM cut of a blade. Their contents are shown in the table below. The designation, "xx," corresponds to a set of data (examples: Base, Cut01, ... Cut06). Both DC and AC-coupled motor data were provided, with a corresponding change in the A/D voltage range. Additionally, two different setups were used to acquire data. Setup 1 contained torsional natural frequencies at 213 Hz and 265 Hz. Setup 2 contained torsional natural frequencies at about 195 Hz and 240 Hz. Blade frequencies are visible in the torsional spectrum within the range of 199-207 Hz.

| Filename | Variable name | Description | Sample rate (Hz) | length of block (N) | A/D voltage range (V) |
|---|---|---|---|---|---|
| xx_motor_dc.mat | | Motor data, DC coupled | | | |
| | fs | sample frequency | | 1 | |
| | Motor_current | Motor current | 12,800 | 1536000 | 5 |
| | Motor_voltage | Motor voltage | 12,800 | 1536000 | 5 |
| | Tach | 1 PPR tach signal | 12,800 | 1536000 | 10 |
| xx_motor_ac.mat | | Motor data, AC coupled | | | |
| | fs | sample frequency | | 1 | |
| | Motor_current | Motor current | 12,800 | 1536000 | 0.05 |
| | Motor_voltage | Motor voltage | 12,800 | 1536000 | 0.01 |
| | Tach | 1 PPR tach signal | 12,800 | 1536000 | 10 |
| xx_accel.mat | | Acceleration data | | | |
| | fs | sample frequency | | 1 | |
| | Vert_accel | vertical acceleration | 12,800 | 1536000 | 1 |
| | Horz_accel | horizontal acceleration | 12,800 | 1536000 | 0.1 |
| | Tach | 1 PPR tach signal | 12,800 | 1536000 | 10 |
| xx_torsion.mat | | Torsional data | | | |
| | Torsion | Digital demodulated torsional vibration signal | 7770.98 | 830139 | na |
| | SampleTimes | sample times of the torsioinal signal | 19.66x10^6 | 830139 | na |
| | fs | average sample frequency of torsional | | 1 | |

Motor voltage, current, and acceleration were supplied as a pure voltage signal. Conversion of these voltages to standard units requires scaling factors as listed in the following table:

| | Conversion Ratio |
|---|---|
| Motor Voltage | 200 Vin/10 Vout |
| Motor Current | $2.5+(0.625*I_P/I_{PN})$ V |
| | $I_P$ = input current |
| | $I_{PN}$ = 6A |
| Vertical Accleration | 1008 mV/g |
| Horizontal Acceleration | 104.4 mV/g |

# APPENDIX D: BEARING PROGNOSTIC TEST RIG

Terri A. Merdes

May 7, 2002

## INTRODUCTION

The Bearing Prognostic Test Rig (BPTR) provides bearing transition-to-failure data that supports development of diagnostic and prognostic algorithms. These algorithms can benefit a wide range of applications and industries, since nearly all rotating machines rely on some type of bearing. With rigorous demands being placed on their loading capacity, bearings are among some of the most important components in rotating machines. There is a trend toward increased reliability requirements. Today rolling bearings have developed into a particular branch of engineering research.

This research enables calculation of bearing life with considerable accuracy, in order to correlate the bearing life with the service life of the machines involved. Unfortunately, a bearing sometimes does not attain its calculated rating life. There may be many reasons for this: inadequate or unsuitable lubrication, heavier loading than has been anticipated, ineffective sealing, careless handling, or insufficient internal bearing clearance. Each fault type produces a particular type of damage, leaving a unique imprint on the bearing itself. Therefore, a growing need exists for algorithms to detect progression of these failures before costly teardown and inspection processes are implemented.

## DEFINING BEARING LIFE

Bearing life is defined as the number of revolutions by the bearing before the first signs of fatigue failure appear. Under normal conditions, the working surfaces of a bearing generally are subjected to very high alternating stresses due to the continuous action of the rolling elements (number of revolutions and magnitude of the load). Specifically, cyclical shear stresses appear immediately below the load-carrying surface, eventually causing cracks which extend out to the outer bearing surface. As the rolling elements pass over these cracks, the surface materials break away in a relatively long, drawn-out process. This fault condition is known as spalling or flaking, which causes increased noise and vibration in the bearing.

## DEEP GROVE BALL BEARINGS

These bearing types are the most widely used of all for general applications, incorporating deep, uninterrupted raceways which makes the bearing suitable for many different loads types: radial loads or radial and axial loads in either direction. Due to the optimum size of the balls and their conformity with raceways, the bearings have a comparatively high load carrying capability and are suitable for high-speed operations and have some misalignment capacity. The bearings are manufactured with evenly spaced balls inside a one or two-piece cage around the raceways, and may be shielded and sealed.

## THE EXPERIMENT

The Bearing Prognostic Test Rig (Fig. D.1) uses a pair of double row spherical roller support bearings to support a shaft on which the test bearing is held in place by a bearing holder between the two support bearings. The test rig was configured to test deep groove ball 1-1/8" test bearings, but is configurable to other types of bearings. A load jack and load cell were mounted directly behind the test bearing and have a radial load capacity of 1,000 lbs in phase with the outer ring. The system is driven by a SCR motor as shown in Table D.1, with an integral tachometer fitted to the motor. The motor is coupled to the test shaft with a flexible coupler to reduce transmitted vibration from the motor. Piezoelectric accelerometers are currently utilized with frequency ranges between 10 and 100KHz. Data is collected on a 48 channel National Instrument 4472 PXI data processing system. The system was run continuously in the

overloaded state, with data acquired at periodic intervals until the bearing finally failed. The objective of this test was acquisition of data throughout the entire test sequence as a quantitative indicator of the failure progression. The deliverables are the data and their historical characterization.

**Table D.1. dc Variable Speed Motor**

| Motor kW | Base Rpm | Arm Vdc | Arm FLAmps | Approx Kgs |
|---|---|---|---|---|
| 1.5 | 3000 | 180 | 10.0 | 31.5 |

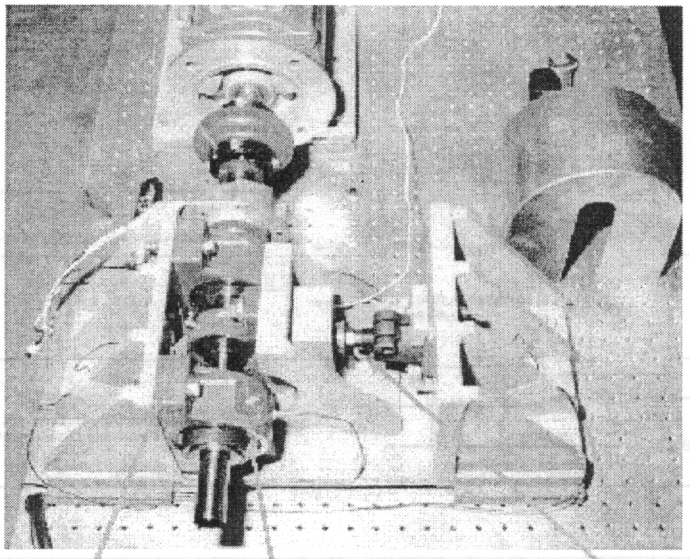



**Figure D.1. Bearing Prognostic Test Rig**

**Table D.2: Tri-axial Accelerometer Description**

| Test Condition | Speed | Radial Load |
|---|---|---|
| 1 | 2500 rpm | 500 lbs |
| 2 | 1250 rpm | 500 lbs |

## CURRENT EQUIPMENT REQUIREMENTS

For this experiment, the BPTR required the purchase of the following:

- A new stainless steel shaft for $65.00, involving machine shop work as shown in Figure D.2.
- Each test deep groove ball bearing from ESI Bearing Distribution costs $5.50.

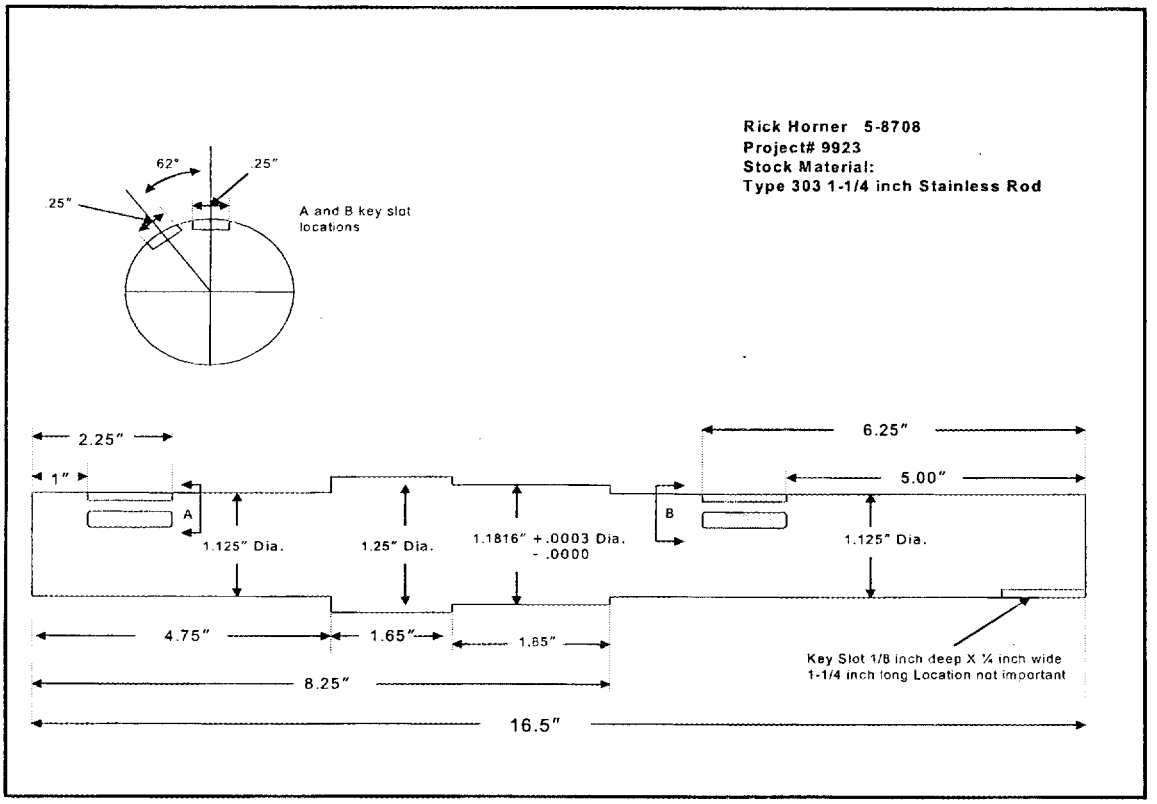



**Figure D.2. BPTR Stainless Steel Shaft**

Two tri-axial accelerometers were installed, as described in Table D.2:

- One on the top center test bearing support housing
- One on the outside roller support bearing

Current and voltage were acquired from the DC electric motor. All data were sampled at 51.2 kHz sample rate using National Instruments 24-bit A/D data acquisition HW (same data acquisition system used on MDTB gearbox tests).

| | |
|---|---|
| **Sensor Directory** | |
| **Sensor Name** | Triaxial Accelerometer |
| **Sensor Make** | PCB |
| | |
| **Sensor Model #** | 356B08 |
| **Sensor Serial #** | 8052 |
| **Sensor Type** | ICP |
| **Sensor Volt Sensitivity** | 100 mV/g |
| | |
| **Bit Resolution** | 24 |
| **Sample Count** | 51200 |

## INTERNAL DISTRIBUTION

1. J. Barhen
2. D. M. Hetrick
3-7. L. M. Hively
8. J. K. Mattingly
9. G. E. Michaels
10. D. L. Moses
11-15. V. A. Protopopescu
16. B. A. Worley
17. T. Zacharia
18. Central Research Library
19. ORNL Laboratory Records – RC
20. ORNL Laboratory Records – OSTI

## EXTERNAL DISTRIBUTION

21. S. R. Martin, DOE, Oak Ridge Operations, P.O. Box 2008, Oak Ridge, TN 37831-6269

22-23. Lynn Hall, U.S. Department of Energy, NE-20/GTN, 1000 Independence Avenue SW, Washington D.C. 20585-1290

24. K. M. Reichard, The Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, PA 16804

25. Ms. Rebecca Richardson, Oakland Operations Office, National Security Programs Department, 1301 Clay Street, Oakland, CA 94612-5208