Prediction of individual cell performance in a large lead/acid peak-shaving battery

Sam P. Perone* and Pei-Hwei Chen**
Department of Chemistry, San Jose State University, San Jose, CA 95192 (USA)

William Spindler
Batteries Included, 15174 Kellen Ct., Riverside, CA 92506 (USA)

(Received July 20, 1992; accepted in revised form December 8, 1992)

Abstract

This work explored the use of routine battery maintenance data for the prediction of individual lead/acid cell performance as measured by capacity values. Pattern recognition techniques were used, and the importance of data-indexing and data-scaling methods for successful prediction was demonstrated. Data-scaling methods addressed the problem of combining or comparing data collected for different batches of cells or for different times. Various data-indexing methods were examined to determine which might capture best the maintenance data trends related to performance prediction. Consistently accurate classification of high- and low-capacity cells verified the existence of information required for performance prediction in the maintenance database.

Introduction

The purpose of the project described here is to examine multivariate relationships among data obtained during maintenance and capacity testing of lead/acid cells used in deep-cycling energy-storage applications, and to evaluate correlations with cell performance. We want to determine if normal maintenance measurements can be used to predict performance of individual cells; to illustrate which maintenance measurements are most informative; and to gain insight to chemical/physical processes affecting performance and life of batteries. The basic approach involves the use of computerized pattern recognition techniques [1–5].

These techniques were evaluated previously [6] for lifetime prediction of individual sealed Ni/Cd cells based on multivariate analysis of manufacturer's fabrication and initial test data. Later [7] it was shown that initial test data for individual lead/acid batteries could be correlated with observed lifetime using multivariate analysis (pattern recognition).

The study described here is directed specifically at predicting performance of individual cells in large energy-storage batteries involving long strings of 100's or 1000's of cells. Unanticipated or undetected cell failures (or cell-reversals due to inadequate

*Author to whom correspondence should be addressed.
**Current address: Industrial Development Bureau, Ministry of Economic Affairs, 41-3, Sinyi Rd. Sec 3, Taipei, 10631 Taiwan, ROC.

0378-7753/93/$6.00

capacity) can cause serious, sometimes catastrophic, damage to battery strings. Our work intends to demonstrate how routine battery maintenance data could provide the predictive ability for early indication of imminent cell failure in large battery installations.

Description of the battery system

This study has focused on maintenance and capacity data collected during operation of a 500 kWh lead/acid energy-storage system operated in a peak-shaving environment by Crescent Electric Membership Corporation (CEMC), Statesville, NC, since July, 1987. The battery consists of 324 2080 Ah cells produced by GNB, Inc. These cells were produced in 1983, following specifications established by the Electric Power Research Institute (EPRI) [8]. The basic performance requirement was to deliver 500 kWh, at 500 kW for 1 h. At this rate the cell-capacity limit was set at 1040 Ah. For a 5 h discharge at the 2080 Ah cell capacity, 1.2 MWh of stored energy could be delivered. The battery was provided with an 8-year warranty.

Fabrication materials and procedures were documented in detail by GNB and associated with individually numbered cells. Cells were produced and formed sequentially in four batches of 80 cells each, and a 5th batch of 20. These series-connected formation groups are referred to as circuits 1 to 5. Fifty-four 6-cell modules were installed at the Battery Energy Storage Test (BEST) Facility (operated by Public Service Electric and Gas Co. for EPRI), and acceptance tests were completed December 7, 1983. The battery underwent over 200 cycles of tests at the BEST Facility from 1983 to 1987.

In July, 1987, the battery was transported to and installed at CEMC, where it has operated as a peak-shaving battery, to discharge at a maximum power of 500 kW for 1 h, or a minimum power of 200 kW for 3 h. Periodic maintenance data, cell impedance measurements, and cell failure observations have been added to the cumulative database. In addition, capacity test data for a carefully selected subset of 109 to 121 of the 340 cells were obtained at CEMC in March, 1989, and April, 1990. Prior to October, 1990, only one cell had been bypassed, due to low capacity. In October, 1990, 44 cells, distributed among 11 modules, were observed to exhibit noticeable case swelling, but have not failed to satisfy capacity requirements.

Statistical, pattern recognition, and cluster analysis studies have been applied to the cumulative database at each stage of the battery's life: after initial fabrication and testing; after cycle testing at BEST, and after several years of operation at CEMC. Results of these studies have been published [9]. They show that variance in fabrication and formation parameters have a profound influence on initial and subsequently measured cell properties. Distinct cell subsets with common fabrication or formation conditions were observed to exhibit similar performance behavior, even into the battery's mid-life period [9].

The work reported here examines maintenance data collected at CEMC in the light of previously-observed relationships between fabrication characteristics and cell subsets with common initial properties. The goal was to determine if subsequent cell performance could be predicted accurately from routine periodic maintenance data. But the earlier cluster analysis observations provided guidance regarding the grouping of cells for statistical and pattern recognition studies based on maintenance data.

Pattern recognition techniques

Pattern recognition involves the detection of regularities among sets of measurements describing objects or events. It includes analytical techniques for processing large amounts of data, the extraction of useful information to reduce the data, and the classification of the data.

A pattern is defined as a d-dimensional vector composed of d independent measurements, and can be represented by:

$$P_i = w_1 x_{1i} + w_2 x_{2i} + w_3 x_{3i} + \dots w_d x_{di} + w_{d+1} \quad (1)$$

where x_{1i}, x_{2i}, $x_{3i} \dots\dots x_{di}$ are variable components (measurements) of the pattern vector for the ith sample, and w_1, w_2, w_3, ... w_{d+1} are constant components (weight vector); d is the number of dimensions.

Because the raw data vector may be of large dimension, some reduction of dimensionality is desired to obtain reliable classification. Thus, a reduced set of N features is extracted from the data which may include combinations and transformations of the raw data (where $N<d$). This reduced feature set should be defined to best characterize the distinguishing properties of each class to which sample i might belong. Numerous systematic techniques have been applied to this task of feature selection, including correlation analysis, statistical distribution analysis, and empirical methods such as sequential iterative feature elimination [1–5, 10, 11].

Pattern classification was done in this work by applying the K–Nearest Neighbor (KNN) classification rule. That is, the class of an unknown item is determined by the majority class of its K nearest neighbors in N-dimensional feature space. K is an odd number, usually one. The inter-item distances in feature space are calculated using Euclidian geometry, where:

$$D_{ij} = \left[\sum_{n=1}^{N} (x_{in} - x_{jn})^2 \right]^{1/2} \quad (2)$$

where i, j are specific items, n the index for all features, and N the total number of features.

The direct visualization of multidimensional feature space is not possible for $N>3$. However, there are several mapping techniques which can reduce the feature space to two dimensions for display. Of these, we have found the nonlinear mapping method [12, 13] to be the most useful. This method transforms N-space data to 2-space data, while attempting to retain the relative inter-item distances which exist in N-space. This is accomplished by an iterative procedure which successfully reduces the 'mapping error' until no perceptible changes occur in the 2-dimensional display. The mapping error is defined as:

$$E = \left[\sum d_{ij} \right]^{-1} \left[\sum (d_{ij} - d_{ij}^*)^2 / d_{ij} \right] \quad (3)$$

where d_{ij} is the distance in N-space and d_{ij}^* the distance in 2-space.

Since the mapping is a 2-dimensional display, it allows visual verification of pattern vector separation by class in N-space. It also allows identification of the nature of classification errors. The nonlinear mapping technique is extremely useful after initial selection of a small set of features, as it guides the selection of optimum feature subsets.

Experimental

Instrumentation

An IBM/AT clone personal computer (PC) was used for database management and computerized pattern recognition. The minimum configuration used included

1 MByte RAM and a 40 MByte hard disk, with a 4 mHz 286 microprocessor. For some applications (nonlinear mapping) a similar system with a 12 MHz 386 microprocessor and 8087 math coprocessor was preferred, because the iterative procedure could be quite time-consuming. A typical pattern recognition run required 8 min of 286 time for 40 patterns and 6 features. A typical nonlinear mapping run for the same data required 210 min of 386 time for 2000 iterations.

Database management and data analysis

All data are contained in a SYMPHONY™ database-management system. Basic statistical computations (averages, variances, distributions, maxima, minima, etc.) and associated graphical procedures were conducted using packages contained within SYMPHONY™. Multivariate analysis procedures (pattern recognition, correlation analysis, and nonlinear mapping) were developed for operation on the PC and programmed in compiled BASIC (Microsoft).

Basic procedures

Feature definition

The raw database used for these studies contained all maintenance data (float voltages, specific gravities, water additions, and electrolyte levels) collected quarterly for all 323 operating cells between August, 1987 and February, 1990. (Subsequent maintenance data are also contained in the database, but were not used for this study.) Because only 109 of these cells underwent capacity testing in both 1989 and 1990, the working database contained only data for these cells.

Features used for pattern recognition studies can be the individual data items (raw data) as well as the transformations or combinations of these data. Thus, in addition to the raw data features, the SYMPHONY database can contain many additional transformed/combined features. In our study, the numbers of features and their relationships to the raw data varied in each different investigation stage. This is because one of the primary goals of our investigation was to determine which transformations would best capture the performance trend information contained in the raw data.

Indexing

Data transformations producing pattern recognition features include: data ratios, trends, sums, differentials, products, etc. In each case, the indexing method has been varied. During our investigation several indexes were considered, of which two are presented here: time sequence of maintenance events, and 'battery-activity' index.

Time-sequence index

This indexing method indicates the sequence of maintenance events during the one-year period preceding a capacity test. Here, we define $(t-1)$ as the time associated with the quarterly maintenance event just prior to a given capacity test; those prior to $(t-1)$ are $(t-2)$, $(t-3)$, and $(t-4)$.

Battery-activity index

The total water added during each quarterly maintenance cycle was considered as an index closely related to battery electrical activity and stress. Thus, the 'water-added' index was defined such that the index $(T-1)$ was assigned to the maintenance event with maximum water addition, $(T-2)$ to the event with the second highest amount of water addition, ..., $(T-4)$ to the event with minimum water addition during the year preceding a periodic capacity test.

Training and prediction sets

The pattern-recognition technique is usually first carried out on a set of known patterns called a *training set* to develop a classifier that recognizes the class membership of these patterns as well as possible; this procedure is called 'training'. The true identity of each pattern is compared with the identity assigned to it in the classification step, and the percentage of correctly-classified training-set patterns is called the *recognition rate*, the classification accuracy of the classifier. If the recognition rate is poor, then other measurements or alternative data transformations should be tried until an acceptably high accuracy is achieved, if possible.

Once a sufficiently accurate decision rule has been obtained by the training procedure, its reliability can be evaluated by observing how successful it is at classifying a different set of known patterns, called a *prediction set*. These patterns have essentially the same origins as the training set and represent the same classes included in the training set. If patterns of the prediction set are also correctly classified with high accuracy, the classification rule can be considered valid and reliable. Truly unknown patterns may then be analyzed as long as each is collected in the same manner as the training/prediction sets and belongs to one of the classes represented in the training/prediction sets.

Because our research goal is to use periodic maintenance data for predicting subsequent cell performance, the maintenance data collected prior to the capacity test in March, 1989, were defined as the training set; and the data preceding the April, 1990, capacity test were defined as the prediction set. The specific feature sets found useful for classifying cells from the training set data were then used (based on the same feature definitions) for prediction.

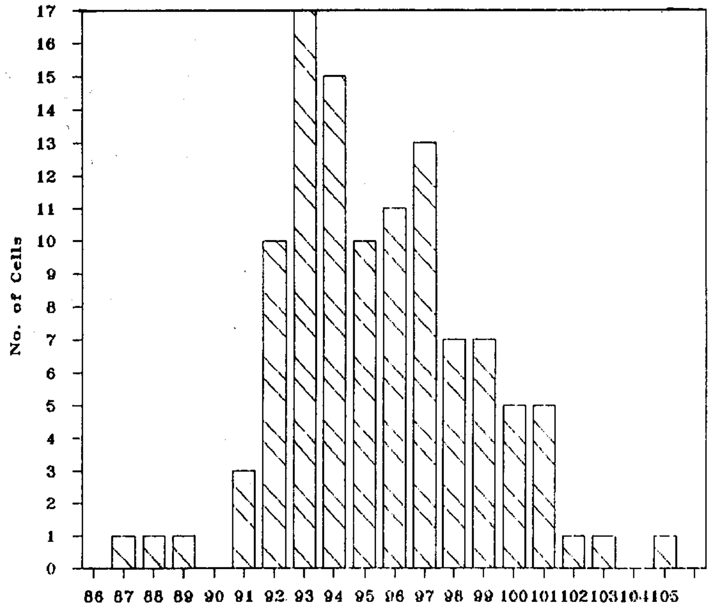
For the CEMC capacity tests, only a selected subset of 109 or 121 of the total 323 cells in operation were monitored for capacity measurements. Thus, although maintenance data were available for all cells, only the 109-cell subset whose capacities were obtained for all tests was considered for defining the training set and prediction set.

Defining class boundaries

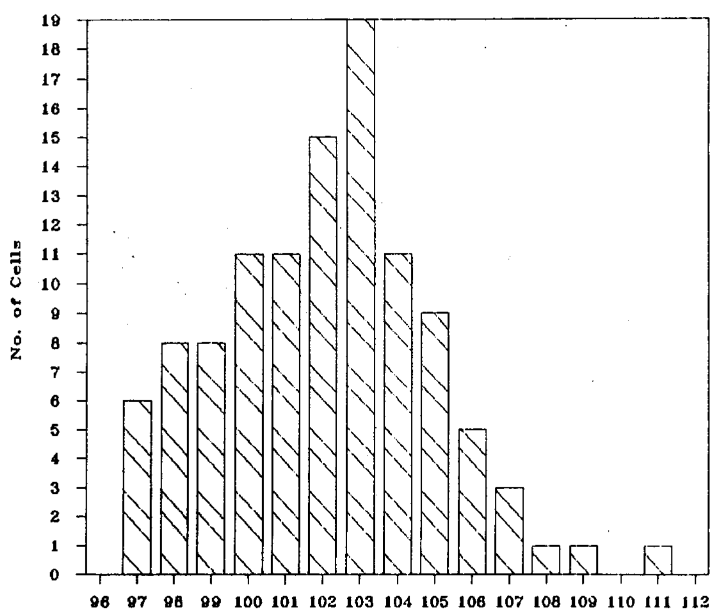
In this study, patterns of both training set and prediction set are divided into three classes according to their capacity values from a specific capacity test. Cells with high capacity are assigned to Class 1, cells with low capacity to Class 2, and intermediates to Class 3. In the computer database all cells were rank ordered according to capacity. We define 'high capacity' cells as those with capacity values $>(\mathrm{AVG}+1\ \mathrm{SD})$, low capacity cells as those with capacity values $<(\mathrm{AVG}-1\ \mathrm{SD})$, intermediates as cells with capacity values within the range $(\mathrm{AVG}\pm 1\ \mathrm{SD})$.

For the training set (maintenance data collected before the March, 1989, capacity test) the average value for 109 cell capacities (measured March, 1989) was 95.0%, and the standard deviation was $\pm 3.1\%$; so Class 1 cells had capacities $>98.1\%$; Class 2 cells had capacities $<91.9\%$; Class 3 cells had capacities between 91.9 and 98.1%. For the prediction set (data collected before the April, 1990, capacity test) the average value for 121 cell capacities (measured April, 1990) was 101.5%, and the standard deviation was $\pm 2.8\%$. Therefore, cells with capacities $>104.3\%$ were assigned to Class 1; cells with capacities $<98.7\%$ were assigned to Class 2; those with capacities in between were assigned to Class 3. Figures 1(a) and 1(b) illustrate the overall distributions of cell capacities in the 1989 and 1990 tests.

The number of cells contained in each training or prediction set varied for each study as discussed below. However, about 65% of the cells belonged to Class 3



(a)



(b)

Fig. 1. Distribution of cell capacities. Discharge current, 450 A, 1.7 V cutoff; capacities corrected to 25 °C. (a) Capacity test at CEMC, March, 1989, 109 cells, and (b) capacity test at CEMC, April, 1990, 121 cells, including all cells from (a).

(intermediate). But only about one-fourth of those cells were included in each study in order to keep the sizes of all classes approximately equal. Selected intermediate class cells were evenly spread in capacity throughout the mid-range.

An alternative procedure for assigning cells to a particular class was to use 'batch normalization'. With this approach the relative capacity values of cells within each fabrication batch (circuits 1 to 5) were considered for assigning classes. The same class boundary criteria were used: i.e., Class 1, capacities >(batch average + 1 SD); Class 2, capacities <(batch average − 1 SD); Class 3, capacities within the range (batch average ±1 SD). Thus, in this case, the absolute values of cell capacities were not compared across batch boundaries.

Normalization of maintenance data

The purpose of normalization is to remove the variance in a data set due to incidental differences in magnitudes of different features. Various normalization methods might be used. For example, conventionally normalized data can be expressed as:

$$(XN_j)_i = (X_j)_i/\bar{X}_j \tag{4}$$

where $(XN_j)_i$ is the normalized data value for the jth feature, ith cell for a specific maintenance event, $(X_j)_i$ the raw data value, and $\bar{X}_j$ the average value of jth feature over all cells for a specific maintenance event.

Trend analysis required normalized data. Trends were obtained by computing a linear regression fit for each cell to the maintenance data over the one-year period preceding each capacity test. Slopes were then used as features for pattern-recognition analysis (see Table 1).

Autoscaling is a more comprehensive normalization procedure which transforms raw data in the following way:

$$(XS_j)_i = ((X_j)_i - \bar{X}_j)/(\mathrm{SD})_j \tag{5}$$

where $(XS_j)_i$ is the autoscaled value for the jth feature, ith cell, for a specific maintenance event, $(X_j)_i$, $\bar{X}_j$, as defined above for eqn. (4), and $(\mathrm{SD})_j$ the standard deviation for the jth feature over all cells for a specific maintenance event.

This autoscaling procedure replaces each data value in a distribution with its distance from the mean expressed in units of standard deviation. All autoscaled distributions have a mean of zero and a standard deviation of one. (This procedure is sometimes referred to as 'standardization'.)

Regardless of prior data transformation, including batch normalization, all pattern features were autoscaled prior to pattern-recognition analysis. Autoscaling allows direct comparison and combination of data sets which were obtained in different time periods. For example, in the pattern recognition *prediction* step, it is necessary to compare training-set and prediction-set data. Based on classifiers developed from training, one prediction set pattern at a time is classified by nearest neighbor analysis with the entire training set. Because the collection conditions might be different for the two data sets, it is reasonable to autoscale the training-set and prediction-set data separately before combining them for pattern recognition.

Results and discussion

The effectiveness of pattern-recognition techniques for battery-performance prediction from maintenance data is critically dependent on data organization procedures.

TABLE 1

Features used for time-sequence index investigations; each item refers to measurements for a specific cell

Feature	Definition
CELVOLT(t-1), ..., CELVOLT(t-4)	Normalized cell float voltage values for (t-1) ... to (t-4), where (t-1) is the maintenance event immediately preceding the capacity test, and (t-4) is the 4th (earliest) event preceding the capacity test
SPGR(t-1), ..., SPGR(t-4)	Normalized specific gravity values indexed as above
LEVEL(t-1), ..., LEVEL(t-4)	Normalized electrolyte level values indexed as above
WATER(t-1), ..., WATER(t-4)	Normalized water addition values indexed as above
NCLVX/Y	[CELVOLT(t-X)–CELVOLT(t-Y)][a]
NSGX/Y	[SPGR(t-X)–SPGR(t-Y)]
NLVLX/Y	[LEVEL(t-X)–LEVEL(t-Y)]
NWATX/Y	[WATER(t-X)–WATER(t-Y)]
AVGVLT	Average value of cell voltages over all maintenance events
AVGSPGR	Average value of specific gravities over all maintenance events
AVGLVL	Average value of electrolyte levels over all maintenance events
AVGWAT	Average value of water-added over all maintenance events
SG * AV	[(AVSPGR) * (AVGVLT)]
AV/SG	[(AVGVLT)/(AVGSPGR)]
PVLT	Product of cell voltages from all maintenance events
RELWAT	Ratio of cumulative water-added (1...4) to average over all cells
SLPCLV	Slope for cell voltages ((t-1)...(t-4))
SLPSG	Slope for specific gravities ((t-1)...(t-4))
SLPLVL	Slope for electrolyte levels ((t-1)...(t-4))
SLPWAT	Slope for water-added ((t-1)...(t-4))

[a] X, Y=1, ..., 4.

Thus, our investigations focused on the effects of various indexing, scaling, and class assignment approaches.

Investigation by time-sequence index

In this part of the study, features were organized by the time-sequence index as described in the Experimental section. Table 1 lists all the features considered in this investigation stage, and defines the indexing nomenclature. All raw data were normalized with respect to the average value of each measurement for each maintenance event (see first four features in Table 1). Normalized data values were used to analyze 'trends', i.e., linear regression slopes were calculated and considered as features for investigation (see last four features in Table 1). Various other features were defined (including differences in normalized values, like NSGX/Y; average values of raw data, like AVGVLT, and other combinations of features, like SG*AV).

Two-class training and prediction

In this study only the high- and low-capacity cells (Classes 1 and 2) were examined to identify the existence of useful features and their discriminating abilities. The training results showed that several feature combinations are capable of discriminating between these two classes of cells with 100% accuracy. When these same features were extracted from the prediction set data, several sets of three to five features gave about 80% overall classification accuracy. These features were: AVGVLT, SPGR(t-2), NCLV3/4, NSG1/3, NLVL2/4, SG*AV, AVGLVL, NLVL3/4, SLPCLV, NCLV1/4.

Three-class training and prediction

The discriminating ability of this feature set (Table 1) for three classes of cell performance (high/low/intermediate) was examined. Three-class training accuracy was only 76–88% overall for the four best feature sets; but Class 1 accuracy was consistently near 100%.

In order to understand the difference in two- and three-class training results, typical nonlinear mappings of two-class and three-class feature space were examined. These are shown in Figs. 2(a) and 2(b). The two-class mapping example of Fig. 2(a) illustrates the complete separation of high- and low-capacity cells in this feature space, which corresponds to the 100% training accuracy obtained with several feature sets. The fuzzy class boundaries seen in Fig. 2(b) however, are consistent with the poorer training results typically obtained for the 3-class set. The observed clustering of Class 1 cells is consistent with the high accuracy of pattern recognition of this class.

When a three-class prediction set was examined using the best sets of training features, meaningful classification was not obtained. Thus, it was concluded that this indexing method and the associated feature definitions were inadequate for realistic classifications of unknown cells. Nevertheless, the accurate training results obtained do confirm that the maintenance features selected do contain information related to performance prediction.

Investigation by battery-activity index

Results with the previous indexing method demonstrated clearly that it was possible to predict battery performance between high- and low-capacity cells, but when intermediate-capacity cells were included in the training and prediction sets, accurate three-class prediction could not be achieved. We believe the problem is that the preceding indexing method does not capture perhaps the most important factor dictating short-term cell performance, and that is battery *usage*. Thus, in this stage of investigation, features were ordered to correspond to the quantity of water-added quarterly in order to provide an index related to battery activity. (See Experimental section). All the features considered in this stage are identical to those in Table 1, except that time (t) is replaced by the index, T, where (T-1) refers to the maintenance event with highest overall water consumption, and (T-4) is the lowest.

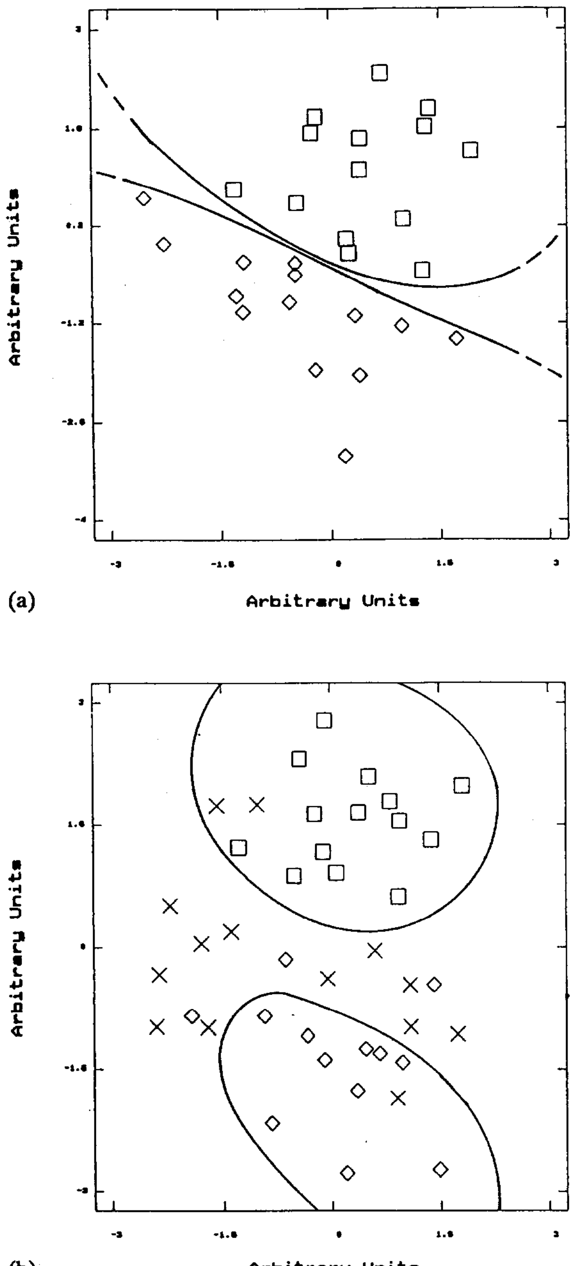


Fig. 2. Nonlinear mapping of cell maintenance features using time-sequence index (see Table 1). Data collected at CEMC prior to March, 1989, capacity test. (□) high-capacity cells (Class 1), (◇) low-capacity cells (Class 2), and (×) intermediate-capacity cells (Class 3). (a) Two-class (high/low capacity) classification, features: NLVL1/2, AVGVLT, AVGSPGR. (b) Three-class (high/low/intermediate capacity) classification, features: SLPWAT, CELVOLT(t-2), CELVOLT(t-1), SG$*$AV.

Two-class training and prediction

Investigation of two-class training again achieved 100% overall classification accuracy with several different feature sets. Results of two-class prediction, using the same feature definitions from successful training, demonstrated that 91 to 94% prediction accuracy could be achieved with as few as two or three features (including AVSPGR, CELVOLT(*T*-2), CELVOLT(*T*-3), NCLV2/3, NCLV3/4). Based on the 100% recognition rates in training and the results of improved and consistent prediction accuracy, we conclude that the water-usage index is the best of several methods tried to organize features to capture the information related to performance prediction.

Three-class training and prediction

Results of three-class training using the water-usage index are shown in Table 2, where the feature sets providing the four best overall classification accuracies are listed. Application of the training-selected features to the 1990 prediction set provided the first meaningful three-class prediction results, and these are listed in Table 3. The same cells as in the training set were used for prediction, but some of them changed their class in 1990.

Although not highly accurate, (57% overall prediction accuracy), some feature sets provided over 90% accuracy in identifying good performing cells (Class 1). Thus, these results were very encouraging and suggested that three-class prediction accuracy may be improved with further refinements. (See below.)

Batch normalization of capacity

The use of relative cell-capacity values and the variance within batches to assign classes was investigated. The same criteria for class assignment boundaries were used (± 1 SD), except that these were applied separately to each of five different batches of cells, where each batch had historically undergone separate fabrication processing and were known to have significantly different average capacities [9]. These batches are referred to as 'circuits 1–5' (see Introduction). This meant that some cells from different batches might be assigned to the same class, despite significantly different capacities, because the average capacities for each batch could be significantly different [9].

After re-assigning cell classes based on batch-normalized cell capacities, pattern-recognition performance prediction studies were done as before. Two-class training resulted in 91 to 94% recognition, while prediction with the same four feature sets resulted in only 64 to 80% accuracy. Not only were these results poorer than without batch normalization, but larger feature sets were required. This indicates that the performance prediction information captured in the maintenance data is more closely related to capacity values normalized over all cells produced, regardless of variance from batch to batch. When three-class studies were done, results also showed degraded accuracy. Thus, for further work, performance classification was done using capacity values normalized over all cells.

Class boundary adjustments

An alternative class boundary adjustment considered was based on the observation that some cells were consistently misclassified, regardless of the specific set of features used. By examining the nonlinear mapping plots of the various feature sets which provided the highest overall training accuracy, it was observed that three fairly well-defined separate clusters of cells did exist, but that a few cells were consistently located within the 'wrong' clusters (i.e., misclassified). To develop an alternative classification

TABLE 2

Water-usage index training results for three-class pattern set; maintenance data collected prior to March 1989, capacity test[a]

Overall classification accuracy (%)	Class 1 classification accuracy (%)	Class 2 classification accuracy (%)	Class 3 classification accuracy (%)	Features used	Feature weight
86	100	78.6	78.6	NCLV1/3, NWAT2/4	2
				PVLT, NSG1/4	1
				LEVEL(*T*-2), NLVL1/2	1
83	100	71.4	78.6	NCLV1/3, NWAT2/4	2
				AVGVLT, NCLV3/4, NSG1/4	1
				AVGLVL, NLVL1/2	1
83	100	71.4	78.6	NCLV1/3	2
				PVLT, AVGVLT, NSG1/4	1
				NWAT2/4, LEVEL(*T*-2), NLVL1/2	1
81	92.9	71.7	78.6	AVGVLT	2
				NCLV1/3, NSG1/4	1
				LEVEL(*T*-2), NLVL1/2	1

[a]Total patterns: 42, Class 1 = high-capacity cells (14), Class 2 = low-capacity cells (14), and Class 3 = intermediate-capacity cells (14).

TABLE 3

Water-usage index prediction results for three-class pattern set; maintenance data collected prior to April, 1990, capacity test[a]

Normalization method used[b]	Overall classification accuracy (%)	Class 1 classification accuracy (%)	Class 2 classification accuracy (%)	Class 3 classification accuracy (%)	Features used[c]	Feature weight
CON	57	75	53.9	47.1	AVGVLT	2
					NCLV1/3, NSG1/4	1
AS	55	83.3	46.2	41.2	LEVEL(*T*-2), NLVL1/2	1
CON	57	91.7	35.7	47.1	NCLV1/3, NWAT2/4	2
					AVGVLT, NCLV3/4, NSG1/4	1
AS	45	83.3	30.8	29.4	AVGLVL, NLVL1/2	1
CON	57	91.7	46.2	41.2	NCLV1/3	2
					PVLT, AVGVLT, LEVEL(*T*-2)	1
AS	55	83.3	46.2	41.2	NWAT2/4, NSG1/4, NLVL1/2	1
CON	50	91.7	30.8	35.3	NCLV1/3, NWAT2/4	2
					PVLT, NSG1/4	1
AS	48	83.3	38.5	29.4	LEVEL(*T*-2), NLVL1/2	1

[a]Total patterns: 42, Class 1 = high-capacity cells (12), Class 2 = low-capacity cells (13), Class 3 = intermediate-capacity cells (17).
[b]CON = conventional normalization, AS = autoscaling normalization.
[c]Features based on Table 1 definitions, except (*T*-*n*) refers to water-added index (see text).

criterion, it was assumed that, when three-class training accuracy $>80\%$ was achieved (based on the ± 1 SD class boundaries), the spatial clustering provided a valid specification of class boundaries. The cluster-generated class boundaries are illustrated in Figs. 3(a) and 3(b), where the KNN misclassified cells are also indicated.

Training was repeated with new class assignments based on the consensus cluster boundaries observed with the six feature sets providing the best training accuracy with the ± 1 SD class boundaries. The results of this subsequent training exercise, of course, provided very accurate recognition of all three classes. More importantly, new feature sets were identified which more sharply separated the three clusters. The training results and feature sets are listed in Table 4.

Prediction results with the feature sets listed in Table 4 were then obtained with the 1990 capacity data, but the ± 1 SD class boundary criterion was used to define prediction set class assignments, as usual. The results are summarized in Table 5. All cells in high- and low-capacity classes were used, but the number of intermediate capacity cells was reduced to keep the size of each class about equal. From Table 5 we can see the overall accuracy for three-class prediction improved (obtaining as high as 67%). Moreover, the individual classification accuracies for the high- and low-capacity cells can be high for different feature sets (93% for Class 1 and 90% for Class 2).

It is important to note that this last training procedure used is based on visually observed clusters in the training set, two of which are *mostly* high- and low-capacity cells, respectively. This approach provides an indirect pathway to our goal of accurately predicting which cells will fall in the high- or low-capacity classes of the prediction set. The fact that the prediction results are good indicates that the clustering (e.g., in Fig. 3) is indeed performance-related.

Observations

Indexing method

The results of our investigations suggest that the 'battery activity' (water-added) index is the best method for maintenance data organization for performance prediction. In fact, it might be preferred to use actual battery-activity data for indexing when it is available. The water-added index is, at best, an indirect measure of the activity history, but it is particularly useful because it may well represent the integrated effects of electrical usage/stress over the preceding maintenance event period.

Because the importance of the 'water-added' method was not recognized prior to our study, no special emphasis was placed on the consistent and accurate collection of those data. Thus, the watering procedures were not always conducted with the purpose of capturing accurately the differences of water required from cell to cell.

Nevertheless, the total water added over all cells during a particular maintenance event can be compared validly with the totals for other maintenance events. In the future, however, the quality of the water-added data should be improved in order to provide more statistical significance to the cell-to-cell variance.

Data scaling methods

The evaluation of data 'normalization' methods was a crucial part of this investigation. For example, in order to compute 'trend' features, raw maintenance data were normalized so that the variance due to changes in absolute data levels from one maintenance event to another would be removed from the computed features. Also, normalization of capacity data within each fabrication batch of cells was investigated to determine if class boundaries should be assigned on a batch-by-batch basis. In the first case, it

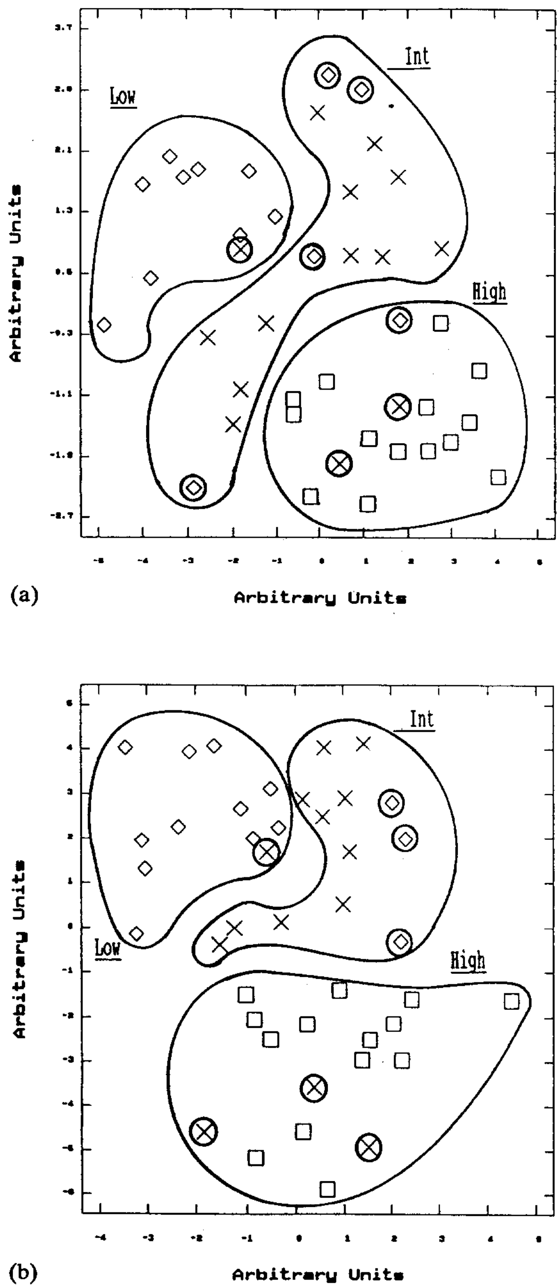


Fig. 3. Nonlinear mapping of cell maintenance features using water-added indexing method. Three-class training (see Table 2). Data collected at CEMC prior to April, 1990, capacity test. (□) high-capacity cells (Class 1), (◇) low-capacity cells (Class 2), and (×) intermediate-capacity cells (Class 3); circled items were misclassified by KNN pattern recognition. (a) 5-dimensional plot, features: AVGVLT, NCLV1/3, NSG1/4, LEVEL(T-2), NLVL1/2. (b) 6-dimensional plot, features: PVLT, NCLV1/3, NSG1/4, NWAT2/4, LEVEL(T-2), NLVL1/2.

TABLE 4

Water-usage index training results for three-class pattern set; class boundary adjustments based on nonlinear mapping; maintenance data collected prior to March, 1989, capacity test[a]

Overall classification accuracy (%)	Class 1 classification accuracy (%)	Class 2 classification accuracy (%)	Class 3 classification accuracy (%)	Features used[b]	Feature weight
100	100	100	100	NCLV2/3	4
				NCLV1/2, SPGR(*T*-1), NSG1/2	1
100	100	100	100	NCLV1/3, NCLV1/4, NCLV2/3	1
				NCLV3/4, NSG2/4, SG * AV	1
				WATER(*T*-1), LEVEL(*T*-2)	1
100	100	100	100	NCLV1/4, NCLV3/4, NSG1/2	1
				NSG2/4, SG * AV	1
				WATER(*T*-1), CELVOLT(*T*-1)	1
97.6	100	90.9	100	CELVOLT(*T*-1), NCLV3/4	1
				SG * AV, NLVL1/4	1

[a]Total patterns: 42, Class 1 = high-capacity cells (14), Class 2 = low-capacity cells (14), and Class 3 = intermediate-capacity cells (14).
[b]Features used based on Table 1 definitions, except (*T*-*n*) refers to water-added index (see text).

TABLE 5

Water-usage index prediction results for three-class pattern set; class boundary adjustments based on nonlinear mapping; maintenance data collected prior to April, 1990, capacity test

Normalization method used[b]	Overall classification accuracy (%)	Class 1 classification accuracy (%)	Class 2 classification accuracy (%)	Class 3 classification accuracy (%)	Features used[c]	Feature weight
CON	58	92.9	57.9	31.6	NCLV2/3	4
					NCLV1/2, NSG1/2, SPGR(*T*-1)	1
AS	60	92.9	63.2	31.6		
CON	54	92.9	68.4	10.5	NCLV1/3, NCLV1/4, NCLV2/3	1
					NCLV3/4, NSG2//4, SG * AV	1
AS	54	92.9	68.4	10.5	WATER(*T*-1), LEVEL(*T*-2)	1
CON	58	64.3	89.5	21.1	NCLV1/4, NCLV3/4	1
					NSG1/2, NSG2/4, SG * AV	1
AS	38	64.3	47.4	10.5	WATER(*T*-1), CELVOLT(*T*-1)	1
CON	67	92.9	78.5	36.8	NLVL1/4	2
					NCLV3/4, SG * AV	1
AS	52	85.7	42.1	36.5	CELVOLT(*T*-1)	1

[a]Total patterns: 52, Class 1 = high-capacity cells (14), Class 2 = low-capacity cells (19), and Class 3 = intermediate-capacity cells (19).
[b]CON = conventional normalization, AS = autoscaling normalization.
[c]Features based on Table 1 definitions, except (*T-n*) refers to water added index (see text).

was found that normalization was essential; whereas, for the second case, normalization by batch was counter-productive.

It was observed that both conventional normalization and autoscaling are important data-conditioning techniques, and that it is crucial to choose the most appropriate technique for a given procedure. For each of the two cases described above, conventional normalization was considered the appropriate scaling technique. However, when computed features were scaled just prior to applying KNN pattern recognition, the autoscaling method was considered the most appropriate for both training and prediction.

When combining training and prediction sets for prediction classification, conventional normalization separately (before combination) usually resulted in higher accuracy prediction compared with combined autoscaled sets. This suggested that the conventional normalization procedure, which preserves the differences in data range, retained more information related to performance prediction. The combined data set was then autoscaled before prediction.

Classification accuracy

The high accuracy of two-class training/prediction studies throughout this investigation certainly verifies the existence of information required for performance prediction in the routine maintenance data. However, three-class prediction is required for practical applications.

Results obtained for three-class training/prediction in this investigation were encouraging, and some observations could lead to practical applications. For example, with some feature sets it was possible to identify either Class 1 (high capacity) cells or Class 2 (low capacity) cells with high accuracy. Thus, if the purpose of a performance-prediction exercise were to preselect a set of high capacity cells for a critical mission, then a feature set providing high Class 1 accuracy with few or no 'false positives' would be very useful. On the other hand, if prior identification is desired of operating cells whose capacity is likely to be unacceptably low in the near future, then high prediction accuracy for Class 2 (low capacity) cells would be very useful. This latter case is ideally suited for the application of pattern recognition to maintenance data, and would be useful even if some 'false positives' were included with the projected low capacity cells.

Examples of feature sets which appear to be useful for either Class 1 or Class 2 prediction are seen in Table 5. For example, the fourth feature set in Table 5 (conventional normalization data) provides 92.9% Class 1 accuracy. The third feature set in Table 5 provides 89.5% accuracy for Class 2 prediction.

Useful features

Although many different features were found useful in various parts of this study, only those features found useful with the water-added indexing method will be discussed here. The most useful features included AVGVLT and those which compute changes in normalized values of CELVOLT, SPGR, WATER, and LEVEL between specific maintenance events. It is not surprising that the voltage changes and the AVGVLT features are useful, as it is expected that the cell float voltage would deteriorate with decreased capacity. It is also significant that those features which document the cell-to-cell variance of *changes* in (normalized) data levels between maintenance events (e.g., NCLV1/3, NSG1/3, etc.) are useful; whereas, those features which document the cell-to-cell variance in (normalized/autoscaled) data levels themselves from one event to another (e.g., CELVOLT(T-X), SPGR(T-X), ... etc.) are not very useful. This observation suggests that a wide range of levels for each of the maintenance parameters are tolerable within each class of cell performance, whereas significant change (or lack of change) in measured levels of maintenance parameters with battery usage, may be indicative of projected performance.

Other available information

This study focused exclusively on the multivariate examination of maintenance data. However, it is important to recognize that other information might also be available to contribute to performance prediction. For example, initial capacity data (at fabrication) for individual cells; capacity data obtained during operation, or initial fabrication/test data might be available for analysis in addition to the routine maintenance data. It would certainly be appropriate to include those additional data for pattern-recognition performance prediction. In fact, we are currently conducting a separate study of the GNB 500 kWh battery initial fabrication/capacity test data for long-term cell performance prediction [14].

The significance of the work reported here, of course, is that in many cases, a battery system will be installed without detailed documentation of fabrication or initial capacity for individual cells, nor will periodic capacity measurements for individual cells be practical in most battery energy-storage systems. Thus, performance prediction of individual cells, at best, will be dependent on examination of routine maintenance data or other readily obtainable periodic observations.

Conclusions

This pattern-recognition investigation clearly demonstrated that performance prediction information is present in maintenance data for flooded lead/acid cells. The key to extracting information for accurate three-class pattern-recognition analysis is to define appropriate methods for indexing data and for minimizing effects of incidental factors affecting data variance. In our investigation the 'battery-activity' index and various data normalization procedures proved to be the effective database treatments.

Although high overall prediction accuracy was not obtained for the three-class problem, selected feature sets do provide high accuracy for identifying either high- or low-capacity cells, with no misclassification between these two classes. Thus, the results of our investigation suggest that it may be possible to use pattern recognition to identify problem cells before failure, and prevent damage to large battery strings. It should also be possible to identify subsets of high-capacity cells: and this should be useful when attempting to place cells in long strings for selected or matched performance. This is potentially as useful as identifying problem cells [15].

Based on this study we have learned much about indexing, normalization, and feature definitions which will benefit future studies on battery performance prediction. It should be pointed out that, although the 'batch normalization' did not facilitate performance prediction of individual cells, it is a significant concept when cells with origins different from training-set items are considered for prediction. Therefore, future studies will need to focus on the feasibility of predictive classification of different sets of uniformly fabricated cells with 'batch-normalized features'. The reliability of feature definitions should also be examined in future studies to recommend better maintenance procedures for battery performance prediction. Finally, the combination of maintenance data with other available data will be examined to determine if higher overall performance prediction accuracy can be achieved.

The ultimate extension of this work to individual cell lifetime prediction from maintenance data will be done when those data become available for the GNB

500 kWh battery. The extension of this work to predictions for newly produced lead/acid batteries is highly desirable. Moreover, we are exploring the extension of the general principles of this work to valve-regulated lead/acid batteries, where considerably different sets of periodic observations on individual cells will need to be studied.

Acknowledgements

The authors wish to thank the Electric Power Research Institute (EPRI) and San Jose State University for support of this work. The authors also thank Al Pivec and Lloyd Weatherman for assistance in planning and directing the capacity tests at CEMC.

References

1 H. C. Andrew, *Introduction to Mathematical Techniques in Pattern Recognition,* Wiley–Interscience, New York, 1972.
2 K. Varmuza, *Pattern Recognition in Chemistry, Lecture Notes in Chemistry,* Vol. 21, Springer, New York, 1980.
3 M. A. Sharaf, D. L. Illman and B. R. Kowalski, *Chemometrics, Chemical Analysis,* Vol. 82, Wiley–Interscience, New York, 1986.
4 D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte and L. Kaufman, *Chemometrics: A Textbook, Data Handling in Science and Technology,* Vol. 2, Elsevier, Amsterdam, 1988.
5 P. C. Jurs and T. L. Isenhour, *Chemical Applications of Pattern Recognition,* Wiley–Interscience, New York, 1975.
6 W. A. Byers and S. P. Perone, *J. Electrochem. Soc., 126* (1979) 720.
7 S. P. Perone and W. C. Spindler, *J. Power Sources, 13* (1984) 23.
8 R. N. McClellan and S. L. Deshpandé, 500 kWh Lead/Acid Battery for Peak-Shaving, Energy Storage, Testing and Evaluation, *EPRI–EM3707,* Oct. 1984.
9 S. P. Perone, R. Petesch, P.-H. Chen, W. C. Spindler and S. L. Deshpandé, *J. Power Sources, 37* (1992) 379.
10 M. A. Pichler and S. P. Perone, *Anal. Chem., 46* (1974) 1070.
11 Q. V. Thomas and S. P. Perone, *Anal. Chem., 49* (1977) 1369.
12 J. W. Sammon, *IEEE Trans. Comput., C-18* (1969) 401.
13 K. Fukunaga, *Introduction to Statistical Pattern Recognition,* Academic Press, New York, 1972, Ch. 10.
14 R. Petesch, *M.S. Thesis,* San Jose State University, 1992.
15 S. P. Perone and P. Symons, *J. Power Sources,* in press.