JOURNAL OF Materials CHEMISTRY

Computer aided design of NiMH electrodes

Kok Liang Heng,[a] Hong Mei Jin,[a] Yi Li[b] and Ping Wu*[a]

[a]*Institute of High Performance Computing, 89B Science Park Drive, #01–05/08 The Rutherford, Singapore Science Park I, Singapore 118261. E-mail: wuping@ihpc.nus.edu.sg*
[b]*Department of Materials Science, National University of Singapore, Lower Kent Ridge Road, Singapore 119260*

Received 10th December 1998

Although a series of applicable and promising compositions of NiMH electrode have been proposed in the literature, it is time and cost intensive by using experimental methods only. In this paper, the effects of La, Ce, Pr and Nd on charge discharge capacity and cycle life of NiMH electrodes are analyzed by applying various artificial intelligence techniques on literature experimental data. This study shows that an increase in the ratios of La/Ce, Ce/Nd and a decrease in Pr are beneficial for both charge discharge capacity and cycle life of rare earth nickel metal hydride electrodes. In particular, a new design of electrode chemistry and its chemical stability are discussed.

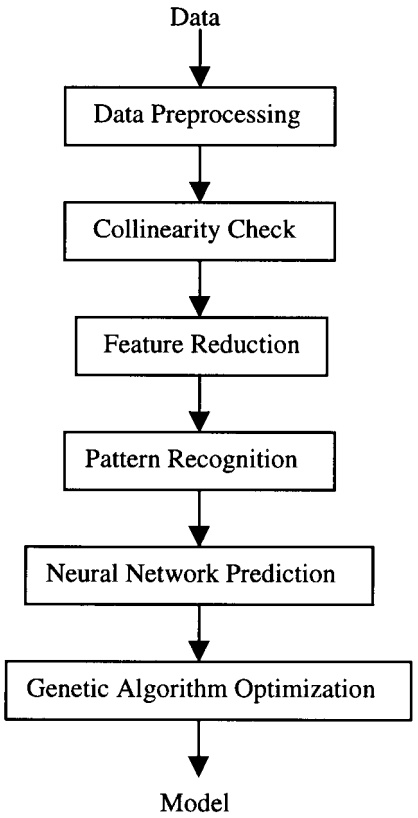
1 Introduction

At the current rate of petroleum consumption the global supply of gasoline will be exhausted in a few decades.[1] Development of alternative energy sources is, therefore, becoming a practical and urgent challenge for scientists and engineers in all countries. Batteries have become one of the most promising alternative energy sources because of their extensive application in our everyday lives.

Different kinds of batteries are available in the market; namely nickel cadmium (NiCd), nickel metal hydride (NiMH) and lithium ion (Li ion). NiMH batteries have many advantages over NiCd batteries with respect to performance and environmental protection, the former can provide up to 30% more capacity than the latter.[2] For Li ion batteries, although they can store on average three times more energy per given weight and twice as much energy per given volume than NiCd batteries,[2] their application is limited by the high costs and concerns on safety and stability issues. Furthermore, NiMH batteries still enjoy a large share in today's battery market.

The aim of artificial intelligence is the development of paradigms or algorithms that require machines to perform tasks that apparently require cognition when performed by humans.[3] APEX (Advanced Process EXpert)[4] is such an artificial intelligence expert system and was developed in house on the IBM SP2 supercomputer of the Institute of High Performance Computing (IHPC). APEX is used as an in-house tool at the Institute of High Performance Computing (IHPC) and will not be available on the market. However, for interest, readers can reproduce the pattern recognition results of this paper, using commercial software such as MATLAB or PARTEK. The tool consists of the following components: data processing, collinearity check, feature reduction, pattern recognition, neural network prediction and genetic algorithm optimization.

Unlike most commercial process diagnosis and optimization techniques, this tool combines the advantages of pattern recognition, neural network and genetic algorithm. It reduces dramatically the effect of noise in the original process data, through a convenient communication between artificial and human intelligence on two dimensional pattern recognition. The model so built is more reliable and accurate. The supercomputing power of IBM SP2 enables a fast solution because the neural network and genetic algorithm operation are very computationally intensive. The package has been successfully applied to the wafer fabrication industry and the petroleum refining industry.[4] Here, APEX will be applied for the design of new NiMH electrodes based on literature experimental data. In recent years, various approaches have been developed for a computer aided materials design.[5–7] However, material design of alloy hydrides tends to be very complex and it is often impossible to do research starting from first principles. The use of APEX reduces the time and cost dramatically. It also demonstrates that APEX can be used in material design application in addition to wafer fabrication and petroleum refining applications. The emphasis of this paper is more on application of the methodology and less about new artificial intelligence methods.



2 Feature reduction

The data used is listed in Table 1. It has a total of 44 records, each record has four predictor variables (La, Ce, Nd and Pr)

Table 1 NiMH electrode data from Guo *et al.*[8] and Jin[9] with four predictor variables (La, Ce, Nd and Pr), two response variables (C_0 and S_{200}) and two class labels (C_1 and C_2)

No.	La	Ce	Nd	Pr	C_0	S_{200}	C_1	C_2
1	0.317	0.135	0.223	0.325	239.4	0.750	2	1
2	0.176	0.206	0.124	0.494	240.5	0.520	2	2
3	0.528	0.029	0.372	0.071	244.7	0.630	2	2
4	0.433	0.040	0.433	0.095	246.4	0.560	2	2
5	0.270	0.135	0.270	0.325	233.7	0.760	2	1
6	0.150	0.206	0.150	0.494	234.3	0.570	2	2
7	0.450	0.029	0.450	0.071	226.9	0.530	2	2
8	0.649	0.040	0.216	0.095	244.2	0.630	2	2
9	0.405	0.135	0.135	0.325	243.4	0.710	2	1
10	0.225	0.206	0.075	0.494	247.2	0.880	2	1
11	0.675	0.029	0.225	0.071	248.1	0.860	2	1
12	0.288	0.040	0.577	0.095	228.0	0.560	2	2
13	0.180	0.135	0.360	0.325	239.8	0.790	2	1
14	0.100	0.206	0.200	0.494	228.8	0.560	2	2
15	0.300	0.029	0.600	0.071	220.8	0.700	2	1
16	0.794	0.066	0.109	0.031	255.6	0.610	1	2
17	0.831	0.069	0.078	0.022	270.7	0.600	1	2
18	0.277	0.023	0.545	0.155	218.3	0.690	2	2
19	0.272	0.272	0.355	0.101	248.2	0.780	2	1
20	0.430	0.430	0.109	0.031	258.0	0.560	1	2
21	0.450	0.450	0.078	0.022	256.3	0.550	1	2
22	0.150	0.150	0.545	0.155	234.1	0.640	2	2
23	0.181	0.360	0.355	0.101	233.2	0.680	2	2
24	0.287	0.573	0.109	0.031	257.7	0.590	1	2
25	0.300	0.600	0.078	0.022	259.2	0.650	1	2
26	0.100	0.200	0.545	0.155	225.6	0.760	2	1
27	0.518	0.026	0.355	0.101	229.5	0.600	2	2
28	0.819	0.041	0.109	0.031	247.8	0.640	2	2
29	0.857	0.043	0.078	0.022	259.0	0.660	1	2
30	0.286	0.014	0.545	0.155	223.5	0.560	2	2
31	0.360	0.450	0.140	0.050	322.0	0.810	1	1
32	0.450	0.050	0.390	0.110	318.0	0.750	1	1
33	0.400	0.350	0.050	0.200	312.0	0.730	1	1
34	0.400	0.200	0.250	0.150	295.0	0.720	1	1
35	0.430	0.430	0.050	0.090	326.0	0.760	1	1
36	0.590	0.060	0.290	0.060	312.0	0.780	1	1
37	0.430	0.300	0.130	0.130	286.0	0.780	1	1
38	0.430	0.300	0.150	0.120	317.0	0.790	1	1
39	0.740	0.050	0.050	0.160	302.0	0.760	1	1
40	0.480	0.240	0.240	0.040	289.0	0.840	1	1
41	0.600	0.170	0.140	0.090	296.0	0.780	1	1
42	0.600	0.050	0.150	0.200	291.0	0.690	1	2
43	0.460	0.270	0.190	0.080	300.0	0.820	1	1
44	0.670	0.230	0.050	0.050	311.0	0.800	1	1

and two response variables (C_0: charge discharge capacity, S_{200}: capacity retention after 200 cycles). For each record, the values of the four predictor variables are summed up to unity. The first 30 records are from Guo *et al.*[8] and the remaining 14 records are from Jin.[9] All records are divided into two classes based on the values of C_0 and S_{200}.

Class label C_1 is assigned to be 1 (good class) if $C_0 \geqslant 250$ and to be 2 (bad class) if $C_0 < 250$. Similarly, class label C_2 is assigned value 1 (good class) if $S_{200} \geqslant 0.7$ and to be 2 (bad class) if $S_{200} < 0.7$.

The data in Table 1 is input to the APEX software. The data is first preprocessed and normalized. This step includes cleaning and scaling of the data. Simple statistics for each variable such as minimum value, maximum value, mean, standard deviation and variance are provided. The data is then scaled so that each variable (predictor and response) has a mean of 0 and a standard deviation of 1. The normalized data then undergoes a collinearity check (CC)[10] and feature reduction by principal component regression (PCR),[11,12] Kruskal–Wallis test (KW)[13] and partial least squares regression (PLSR).[14] The sample correlation matrix is calculated in the CC module. The correlation coefficient is checked for each pair of predictor variables. CC is considered as a preliminary step to screen out variables that are collinear.

To further eliminate variables that are not important, the feature reduction module is used. It consists of PCR, KW and PLSR. Detailed information about these algorithms can be found in the respective references. Here the algorithms will be presented very briefly. In PCR, the matrix $\boldsymbol{X}$ containing the predictor variables is expressed as a score matrix $\boldsymbol{T}$. This matrix $\boldsymbol{T}$ results from the projection of $\boldsymbol{X}$ onto the space defined by the eigenvectors $\boldsymbol{W}$ as

$$\boldsymbol{T} = \boldsymbol{X}\boldsymbol{W}$$

The first principal component of the matrix $\boldsymbol{T}$ is simply the linear combination of the original variables that has the smallest error when used to estimate the original variables. The second step of the PCR is to regress the matrix $\boldsymbol{Y}$ containing the response variables onto the score matrix as

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{B} + \boldsymbol{E}$$

$$B = (T^{*}T)^{-1}T^{*}Y$$

where * denotes matrix transportation and $\boldsymbol{E}$ is the error matrix. If all the eigenvectors are used to form $\boldsymbol{T}$, PCR will yield results identical to multivariate linear regression. The advantage of PCR over multivariate linear regression is that PCR concentrates useful information into fewer factors (here factors mean some linear combination of the original variables) and the other factors containing more noise than information can be deleted by cross validation[15] and calculation of PRESS (prediction residual error sum of squares).[15] The method PLSR is a modeling procedure that estimates underlying factors in $\boldsymbol{X}$ and $\boldsymbol{Y}$ simultaneously. The approach taken by PLSR is very similar to that of PCR, except that the factors are chosen to describe the variables in $\boldsymbol{X}$ as well as in $\boldsymbol{Y}$. Similarly, cross validation[15] and PRESS estimation[15] can be used to determine the number of factors needed to model $\boldsymbol{Y}$ by $\boldsymbol{X}$. KW is designed to test that if a variable is important in the classification of patterns. The test is independent of the response variable.

The feature reduction module involves variable scoring and variable selection. In variable scoring, predictor variables are scored based on their relative importance as returned by PCR, KW and PLSR. In case of KW, the relative importance is the KW value. The higher the KW value, the more important the predictor variable is. In the case of PCR and PLSR, their relative importance is defined in term of the absolute regression coefficients of the original predictor variables. The larger the coefficient, the more important the associated predictor variable is. For PCR, the group scoring model will equally divide candidates in the list returned by PCR into three groups. A weight of 1, 2 or 3 is assigned to the groups according to their importance factors as judged by PCR. The most important factors will be assigned the weight value 3. The same treatment applies to both KW and PLSR. In variable selection, an aggregate score is computed for each of the variables. There are three selection models available; namely 1 norm, 2 norm and infinity norm. Variables with aggregate score less than a preset score will be removed. The norms used are defined as follows:

$$\text{1 norm:} \quad \text{agg\_score} = |\text{s1}| + |\text{s2}| + |\text{s3}|$$

$$\text{2 norm:} \quad \text{agg\_score} = (\text{s1}^2 + \text{s2}^2 + \text{s3}^2)^{1/2}$$

$$\text{infinity norm:} \quad \text{agg\_score} = \max(\text{s1}, \text{s2}, \text{s3})$$

where agg_score denotes aggregate score, s1, s2 and s3 are PCR score, KW score and PLSR score returned in the variable scoring step.

PCR, KW and PLSR are all linear methods and they have different criteria and objectives. By adopting the above voting scheme, we reduce the danger of deleting some important variables because variables may be important to one method but not the other two methods. After the feature reduction step, the less important variables will be deleted but the most

significant variables will be kept. Since all the four predictor variables are important to both the targets C_0 and S_{200}, none of them is deleted in the feature reduction module.

3 Pattern recognition

Four important predictor variables remain after feature reduction. In order to visualize pattern and to identify optimal zone, 2D projections of the four-dimensional data are necessary. In the APEX tool, two dimensional projections of the four-dimensional sample space are obtained using various pattern recognition techniques including principal component analysis (PCA),[11,12] partial least squares (PLS),[14] modified Fisher discriminant analysis (MF)[16,17] and linear mapping (LM).[18] Interested readers can read the respective references of these algorithms. A genetic algorithm which will be described in detail later, is used to select a few good projections among all the projections generated by the above four methods. There are two types of separators used in separating good class from bad class, namely line separator and sphere separator. The projection whose error value is the smallest will be the best projection. The error value is determined from the error function relevant to the particular separator. The error functions for the two different separators are detailed below.

$$\text{Line: } E1+E2+|E1-E2|/(E1+E2+1)$$

$$\text{Sphere: } E1+E2+|E1-E2|/(E1+E2+1)+r/R$$

where $E1$ and $E2$ are accumulated values of the number of wrongly classified good class and bad class data points respectively for using the separation function, r is the radius of the sphere and R is the range of the radius searched for in determining the sphere separator. This automatic selection reduces human involvement in choosing the good projections. It is one of the advanced features of APEX and is particularly useful when the number of features is large.

In this paper, PLS is found to give good separation for the target C_0 (Fig. 1). The X- and Y-axis are given by the following equations:

$$X=-0.12+2.43\text{La}+2.43\text{Ce}-4.09\text{Nd}-2.58\text{Pr} \quad (1)$$

$$Y=0.80-2.88\text{La}+3.28\text{Ce}-2.71\text{Nd}+3.12\text{Pr} \quad (2)$$

Substitution of the contents of La, Ce, Nd and Pr into eqn. (1) and (2) results in a point in two-dimensional space. From Fig. 1, it is observed that the two classes are well separated.

Fig. 2 shows a classification of S_{200} according to whether S_{200} is greater or less than 70% by PLS. A high reading in S_{200} means a long cycle lifetime. It can be seen from Fig. 2 that data with S_{200} are distributed in a band-like region

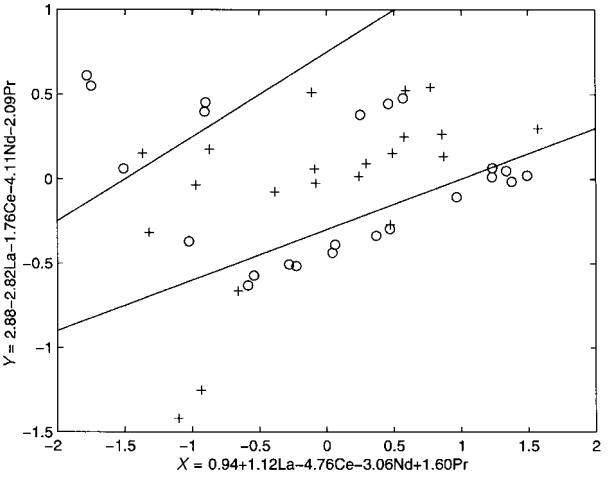


Fig. 2 Classification of samples with different capacity retention after 200 cycles [(+) $S_{200}\geqslant 0.7$, (○) $S_{200}<0.7$].

between two regions of the data with $S_{200}<70\%$. The X- and Y-axis are given by the following equations:

$$X=0.94+1.12\text{La}-4.76\text{Ce}-3.06\text{Nd}+1.60\text{Pr} \quad (3)$$

$$Y=2.88-2.82\text{La}-1.76\text{Ce}-4.11\text{Nd}-2.09\text{Pr} \quad (4)$$

Next, eight new data points from Jiang *et al.*[19] are used to check the above two patterns for C_0 and S_{200} and data is listed in Table 2. All the new data points have $S_{200}>0.7$ and they are all located in the band-like region between the two regions with $S_{200}<0.7$ (Fig. 3). Only two of the eight new data points have $C_0>250$. However, these two points are located at the left hand region of Fig. 4 which is not in agreement with where we believe the optimal region should be (right hand region of Fig. 4). This may be due to experiment

Table 2 New test data from Jiang *et al.*[19] with four predictor variables (La, Ce, Nd and Pr), two response variables (C_0 and S_{200}) and two class labels (C_1 and C_2)

No.	La	Ce	Nd	Pr	C_0	S_{200}	C_1	C_2
1	0.400	0.100	0.400	0.100	232.0	0.844	2	1
2	0.400	0.100	0.300	0.200	245.0	0.856	2	1
3	0.400	0.100	0.200	0.300	290.0	0.835	1	1
4	0.200	0.300	0.400	0.100	210.0	0.828	2	1
5	0.300	0.200	0.400	0.100	240.0	0.793	2	1
6	0.400	0.200	0.300	0.100	258.0	0.802	1	1
7	0.400	0.300	0.200	0.100	220.0	0.832	2	1
8	0.400	0.400	0.100	0.100	205.0	0.810	2	1

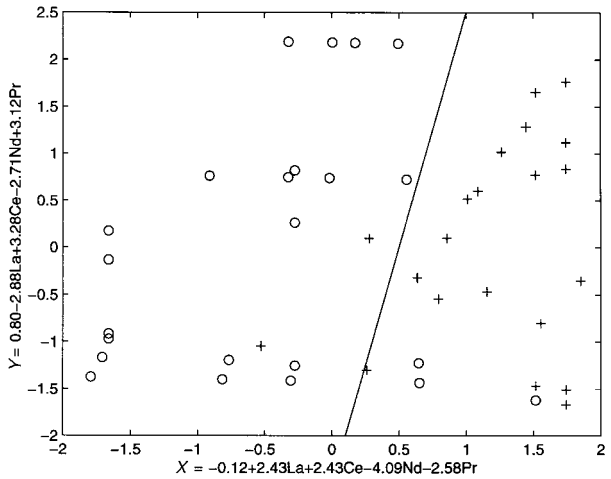


Fig. 1 Classification of samples with different charge discharge capacities [(+) $C_0\geqslant 250$, (○) $C_0<250$].

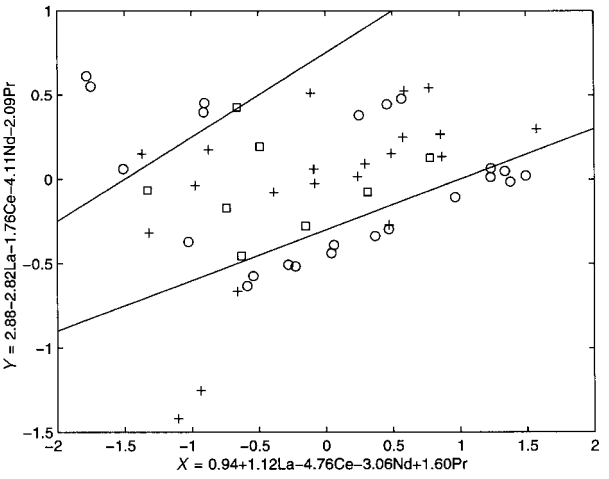


Fig. 3 Location of new test samples using eqn. (3) and (4) [(+) original $S_{200}\geqslant 0.7$, (○) original $S_{200}<0.7$, (□) test $S_{200}\geqslant 0.7$].

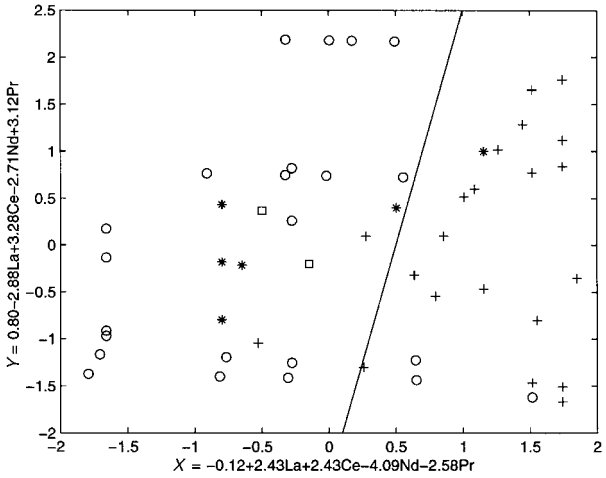


Fig. 4 Location of new test samples using equations (1) and (2) [(+) original $C_0 \geqslant 250$, (○) original $C_0 < 250$, (□) test $C_0 \geqslant 250$, (∗) test $C_0 < 250$].

errors. Another possible explanation is that the electrode studied by Jiang *et al.*[19] is $RE(NiCoMnTi)_5$ while the electrode studied by Guo *et al.*[8] and Jin[9] is $RE(NiCoMnAl)_5$. For the remaining six points which have $C_0 < 250$, only one point is at the wrong side of the boundary, the other five points being located correctly at the left hand region of Fig. 4.

4 Neural network prediction

In order to predict C_0 and S_{200} from the four predictor variables, a neural network is developed for each of the response variables. Because the four predictor variables satisfy the constraint La + Ce + Nd + Pr = 1, it is sufficient to train the network with three of the predictor variables. In this paper, La, Ce and Nd are chosen to train the neural network.

In the APEX software, neural networks are trained using Levenberg–Marquart optimization.[20] Before each network is trained, the input data set is randomly partitioned into a training set, a validation set and optionally a test set. The validation set is used to implement 'early stopping' to help prevent overfitting and to improve generalization and predictability. During the training, the network error is calculated after each epoch (a pass through the training set) using the validation set, and the model with the smallest validation error is retained. The test set, if present, is used to estimate the generalization error for the model that is retained. This module is designed to build and test multiple neural network models efficiently and automatically. It takes advantage of multiple nodes on the IBM SP2 supercomputer using a simple task oriented form of parallelism.

In our case study, a random validation set containing 10% of the samples (four records) is held out to implement 'early stopping' for both the C_0 and S_{200}. The values of C_0 and S_{200} from experiments *vs.* those predicted by the neural networks are shown in Fig. 5–8. The results of predicted values are in agreement with the experimental values.

5 Design of new NiMH electrodes

In the design of new NiMH electrodes, the composition of the four elements (La, Ce, Nd and Pr) is changed and the effect of such a change on C_0 and S_{200} is analyzed. In order to design a new NiMH electrode, economical factors and performance are the selection criteria. Since Nd is the most expensive element among the four, the objective is to design the most effective electrode with as little Nd as possible.

In our first design, the value of Nd ranges from 0.05 to 0.32 in a step increment of 0.03. The ratio of Ce/Nd is kept at 0.5,

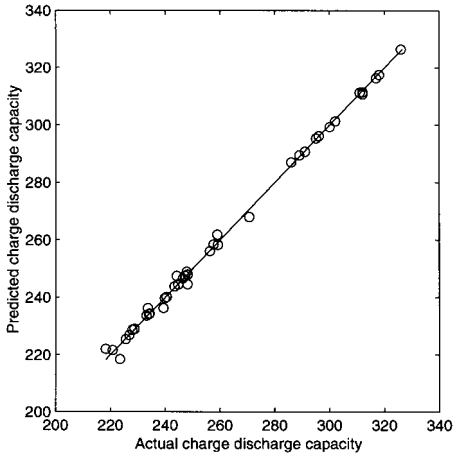


Fig. 5 Actual and predicted values of C_0 (Training set).

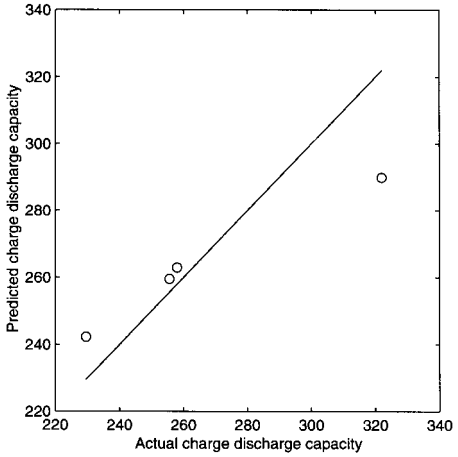


Fig. 6 Actual and predicted values of C_0 (Validation set).

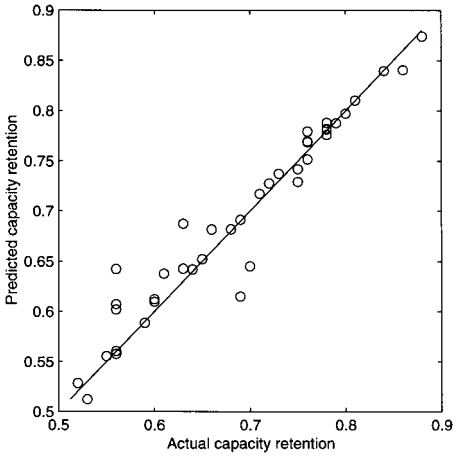


Fig. 7 Actual and predicted values of S_{200} (Training set).

1 or 2 while the value of Pr is kept at 0.05. Fig. 9 gives the performance of S_{200} with respect to these designed parameters. From Fig. 9, it is observed that in order to make the new designed points close to the optimal zone (the band-like region), the ratio Ce/Nd is preferred to be either 1 or 2. But since it is desirable to have as little usage of Nd as possible, the ratio Ce/Nd is perferred to be 2. In a similar design, no significant improvement is found if the value of Pr is set to be 0.15.

The design criteria mentioned above for S_{200} is applied to C_0 and the performance of C_0 with respect to these design criteria is depicted in Fig. 10. Similar to S_{200}, a better result

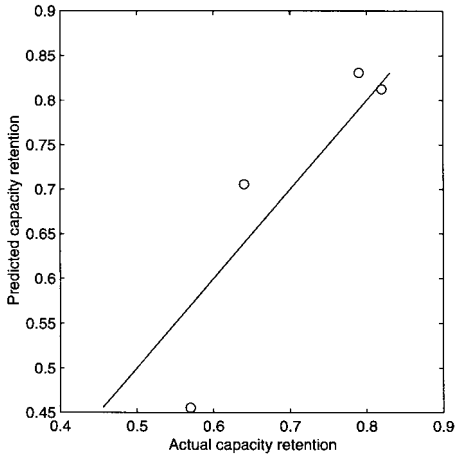


Fig. 8 Actual and predicted values of S_{200} (Validation set).

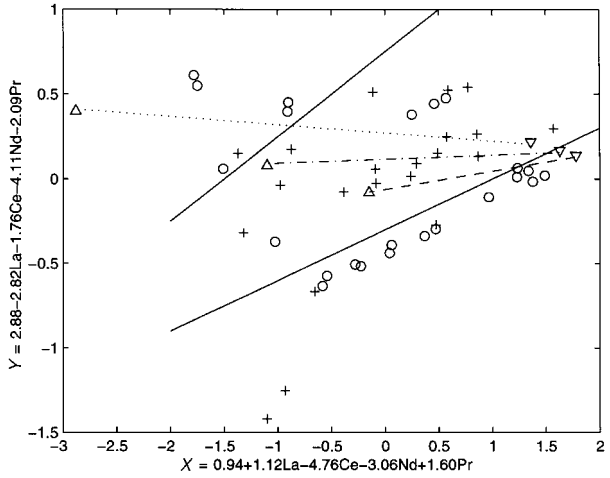


Fig. 9 Performance of S_{200} with respect to Ce/Nd, Pr = 0.05, Nd = 0.05, 0.32, 0.03 [(▽) Nd = 0.05, (△) Nd = 0.32; (···) Ce/Nd = 2, Pr = 0.05; (–·–·) Ce/Nd = 1, Pr = 0.05; (–––) Ce/Nd = 0.5, Pr = 0.05].

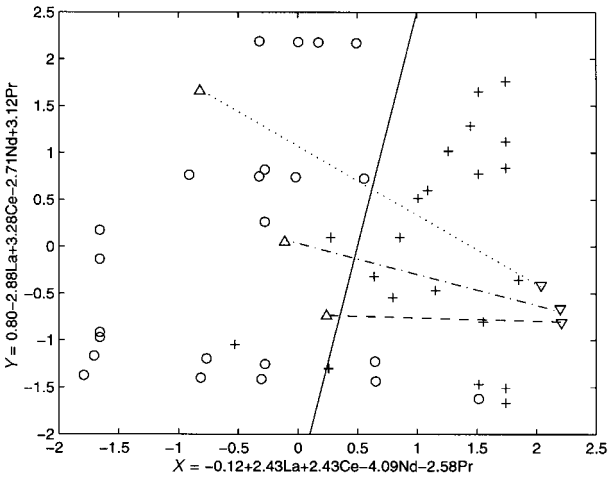


Fig. 10 Performance of C_0 with respect to Ce/Nd, Pr = 0.05, Nd = 0.05, 0.32, 0.03 [(▽) Nd = 0.05, (△) Nd = 0.32; (···) Ce/Nd = 2, Pr = 0.05; (–·–·) Ce/Nd = 1, Pr = 0.05; (–––) Ce/Nd = 0.5, Pr = 0.05].

is achieved if the ratio of Ce/Nd is chosen to be 2. It can be seen from Fig. 10 that the points when Ce/Nd equals 2 are closer to the optimal region than the points when Ce/Nd equals 0.5 or 1. At the same time, the value of Nd should be kept small since a high value of Nd will cause the resultant points to move towards the bad zone (left hand region of Fig. 10). The same conclusion can be drawn if Pr equals 0.15.

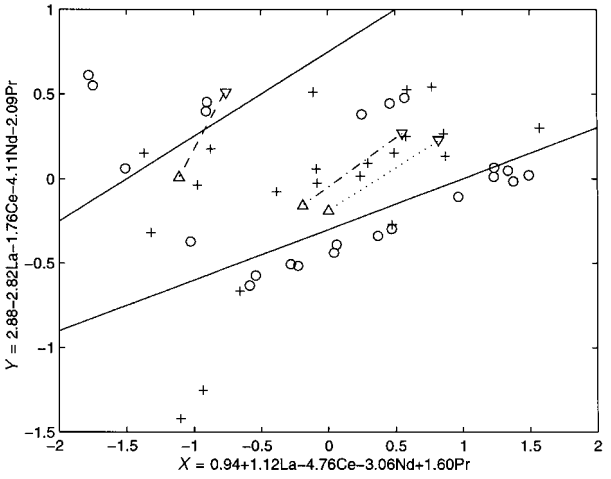


Fig. 11 Performance of S_{200} with respect to La/Ce, Pr = 0.05, Nd = 0.05, 0.32, 0.03 [(▽) Nd = 0.32, (△) Nd = 0.05; (···) La/Ce = 4, Pr = 0.05; (–·–·) La/Ce = 3, Pr = 0.05; (–––) La/Ce = 1, Pr = 0.05].

Next, the effects of the ratio La/Ce on S_{200} and C_0 are investigated. In this case, the ratio of La/Ce is selected to be 1, 3 or 4. The values of Pr and Nd are the same as those used in the first design. Fig. 11 shows the performance of S_{200} while Fig. 12 gives the performance of C_0. It can be observed from Fig. 11 that new points will fall in the optimal zone if the ratio of La/Ce equals 3. The trend is not changed when the value of Pr is changed to 0.15.

It should be noted that C_0 is insensitive to the ratio of La/Ce. Designed points with the ratio La/Ce equal to 1, 3 or 4 are overlapped on a straight line (Fig. 12).

However, the value of Nd should be small as this is the optimal direction to follow (right region of Fig. 12). C_0 remains insensitive to the ratio La/Ce when Pr is set to be 0.15.

The following last set of parameters is used: the ratio of La/Ce can either be 1, 3 or 4, the value of Pr is increased from 0.05 to 0.32 with a step size of 0.03 and the ratio of Ce/Nd is chosen to be 2. For S_{200}, the set of parameters La/Ce = 3, Ce/Nd = 2 gives the best performance (Fig. 13). A low value of Pr should also be used since higher Pr will not bring in any additional benefits.

As for C_0, it is clear from Fig. 14 that La/Ce = 3 and Ce/Nd = 2 give the best result among all the design criteria. A high value of Pr will move the designed points toward the bad zone located at the left region of Fig. 14.

The effects of the ratios La/Ce and Ce/Nd and the value of

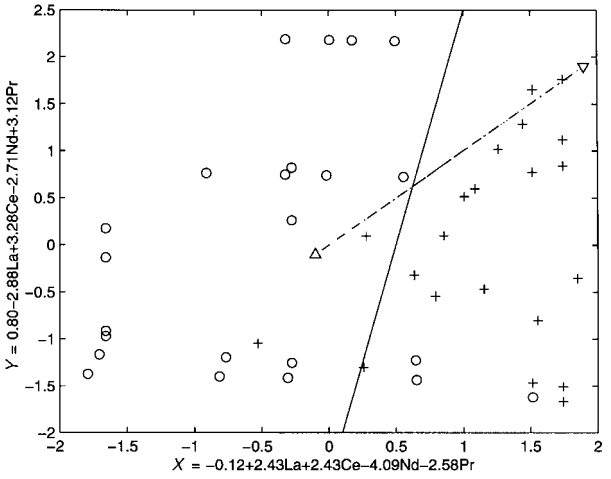


Fig. 12 Performance of C_0 with respect to La/Ce, Pr = 0.05, Nd = 0.05, 0.32, 0.03 [(▽) Nd = 0.32, (△) Nd = 0.05; (···) La/Ce = 4, Pr = 0.05; (–·–·) La/Ce = 3, Pr = 0.05; (–––) La/Ce = 1, Pr = 0.05].

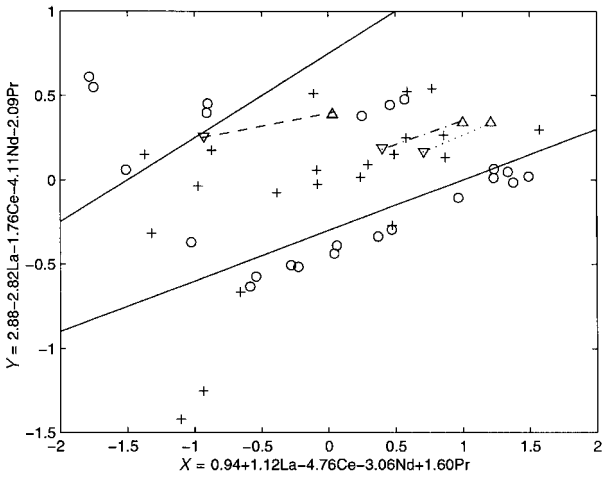


Fig. 13 Performance of S_{200} with respect to La/Ce, Ce/Nd = 2, Pr = 0.05, 0.32, 0.03 [(▽) Pr = 0.05, (△) Pr = 0.32; (···) La/Ce = 4, Ce/Nd = 2; (– · –) La/Ce = 3, Ce/Nd = 2; (–––) La/Ce = 1, Ce/Nd = 2].

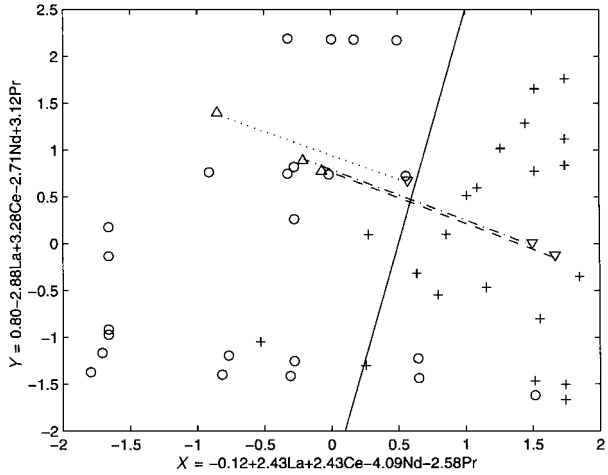


Fig. 14 Performance of C_0 with respect to La/Ce, Ce/Nd = 2, Pr = 0.05, 0.32, 0.03 [(▽) Pr = 0.05, (△) Pr = 0.32; (···) La/Ce = 4, Ce/Nd = 2; (– · –) La/Ce = 3, Ce/Nd = 2; (–––) La/Ce = 1, Ce/Nd = 2].

Pr on the performances of S_{200} and C_0 are studied and analyzed. The study shows that in order to reduce costs and improve performances of S_{200} and C_0, the following criteria can be adopted.

$$\text{La/Ce} = 3 \quad (5)$$

$$\text{Ce/Nd} = 2 \quad (6)$$

$$\text{Pr} = 0.05 \quad (7)$$

Based on the above criteria, the values for the four elements are La = 0.63, Ce = 0.21, Nd = 0.11 and Pr = 0.05. Using these values and the neural network models for S_{200} and C_0, the predicted values of S_{200} and C_0 are 0.8208 and 289.1274 respectively which are considered optimal points for both response variables.

6 Genetic algorithm optimization

A new design is useful only if the requested precision of chemical composition is low enough to match the experimental error limits. Otherwise, it would be very difficult for quality control in the later production process. For the new designed electrode [eqn. (5), (6) and (7)], its performance stability within the range of La = 0.62–0.64, Ce = 0.19–0.23 and Nd = 0.09–0.11 is studied. For this purpose, a genetic algorithm may be used since it is particularly useful for searching the input space for optimal points.

Table 3 Results of maximization and minimization of optimal points using a genetic algorithm

Maximum or minimum	La	Ce	Nd	Predicted C_0	Predicted S_{200}
301.0283 (Maximum C_0)	0.633	0.195	0.121	301.0053	0.8084
271.2917 (Minimum C_0)	0.626	0.193	0.091	271.5224	0.8898
0.884 (Maximum S_{200})	0.626	0.193	0.094	274.1538	0.8832
0.757 (Minimum S_{200})	0.633	0.219	0.117	286.9966	0.7586

A genetic algorithm is a straightforward computerized search method based on the ideas of genetics and natural selection. It relies on the assumption embedded in the idea of natural selection that as members of the population mate, they produce offspring that have a significant chance of retaining the desirable characteristics of their parent, perhaps even combining the best characteristics of both parents. In this manner, the overall fitness of the population can potentially increase from generation to generation. After generations of evolution, an optimal solution will finally be obtained. Genetic algorithms have been shown to solve linear and nonlinear problems by exploring all regions of the state space and exponentially exploiting promising areas through mutation, crossover and selection operations applied to individuals in the population.[21]

In the APEX software, a trained neural network model, which is a function returning an appropriate target value for each set of input values for predictors, will be taken as input to the genetic algorithm module. The genetic algorithm will then perform minimization, maximization or set point optimization on the specified function in which the specified function will be optimized to as close to the stated set point value as possible.

Recall that the neural network models for S_{200} and C_0 are based on the three predictor variables La, Ce and Nd and the optimal values for La, Ce and Nd are 0.63, 0.21 and 0.11, respectively. The variation ranges of La, Ce and Nd are La (0.62–0.64), Ce (0.19–0.23) and Nd (0.09–0.11). The results of maximization and minimization are shown in Table 3. The first column of Table 3 gives the maximum or minimum value of C_0 or S_{200} returned by the genetic algorithm. The second to fourth columns are the values of La, Ce and Nd returned by the maximization or minimization procedure. The predicted values of C_0 and S_{200} by neural network are in the last two columns. From Table 3, the maximum values of S_{200} and C_0 within the studied region are 0.884 and 301.0283 respectively; the minimum values of S_{200} and C_0 within the studied region are 0.757 and 271.2917 respectively. Therefore, the model predicted that within the composition range of La = 0.62–0.64, Ce = 0.19–0.23 and Nd = 0.09–0.13, all electrodes will have at least 0.757 and 271.2917 for S_{200} and C_0 which satisfy the good class criteria for both response variables ($S_{200} \geqslant 0.7$, $C_0 \geqslant 250$).

7 Conclusion

Various artificial intelligence techniques are used in the design of new nickel metal hydride electrodes. It is shown that an increase in the ratios of La/Ce and Ce/Nd and a decrease in the value of Pr improve the charge discharge capacity and cycle life of rare earth nickel metal hydride electrodes. A new electrode is thus proposed:

La = 0.62–0.64

Ce = 0.19–0.23

Nd = 0.09–0.13

A neural network and genetic algorithm are combined together to provide estimates of design values for charge discharge capacity and cycle life within the proposed range.

Acknowledgements

The authors would like to thank Dr. Robert Straughan and Ms Poh Kiang Chua of Institute of High Performance Computing, who contributed to the Genetic Algorithm Optimization module of the APEX tool.

References

1 L. F. Ivanhoe, *Futurist*, 1997, **31**, 20.
2 I. Buchmann, *Batteries in a Portable World, A Handbook on Rechargeable Batteries for Non-Engineers*, ed. S. Gross, J. Lackner and M. Sulkes, Canada, 1997, p. 4.
3 *Concise Encyclopedia of Information Processing in Systems and Organizations*, ed. A. P. Sage, Pergamon, New York, 1990, p. 9.
4 P. Wu, R. Straughan, I. Ong and K. L. Heng, *Australasia-Pacific Forum on Intelligence Processing and Manufacturing of Materials*, ed. T. Chandra, S. R. Leclair, J. A. Meech, B. Verma, M. Smith and B. Balachandra, IPMM'97, p. 235.
5 D. Wienke, K. Danzer, M. Gitter, J. Aures, H. U. Munch, H. G. Byhan and H. J. Pohl, *Anal. Chim. Acta*, 1989, **223**, 247.
6 H. L. Liu, Y. Chen and N. Y. Chen, *J. Chemometr.*, 1994, **8**, 439.
7 H. L. Liu, S. P. Huang and N. Y. Chen, *Patt. Recog. Arti.*, 1993, **6**, 399.
8 J. Guo, C. H. Li, H. L. Liu and N. Y. Chen, *J. Electrochem. Soc.*, 1997, **14**, 2276.
9 H. M. Jin, PhD Thesis, General Research Institute for Non Ferrous Metals, Beijing, 1997.
10 A. A. Afifi and V. Clark, *Computer Aided Multivariate Analysis*, Van Nostrand Reinhold, New York, 2nd edn., 1990, p. 153.
11 J. D. Jobson, *Applied Multivariate Data Analysis*, Springer-Verlag, New York, 1992, vol. 2, p. 346.
12 F. J. Bryan, *Multivariate Statistical Methods*: *A Primer*, Chapman & Hall, London, 2nd edn., 1994, p. 76.
13 J. S. Milton and J. C. Arnold, *Introduction to Probability and Statistics*: *Principles and Applications for Engineering and the Computing Sciences*, McGraw-Hill, New York, 2nd edn., 1990, p. 508.
14 P. Geladi and B. R. Kowalski, *Anal. Chim. Acta*, 1986, **185**, 1.
15 R. H. Myers, *Classical and Modern Regression with Application*, Duxbury Press, MA, 1986, p. 103.
16 G. T. Rasmussen, G. L. Ritter, S. R. Lowry and T. L. Isenhour, *J. Chem. Inf. Comput. Sci.*, 1979, **19**, 255.
17 R. J. Schalkoff, *Pattern Recognition*: *Statistical*, *Structural and Neural Approaches*, John Wiley & Sons, New York, 1992, p. 90.
18 K. Fukunage, *Introduction to Statistical Pattern Recognition*, Academic Press, Boston, 1990, p. 467.
19 J. J. Jiang, Y. Q. Lei, D. L. Sun, C. S. Wang, J. Wu and Q. D. Wang, *Acta Metall. Sinica*, 1985, **31**, 379.
20 M. T. Hagan and M. Menhaj, *IEEE Trans. Neur. Netw.*, 1994, **5**, 989.
21 Z. Michalewicz, *Genetic Algorithms + Data Structure = Evolution Programs*, AI Series, Springer-Verlag, New York, 1992, p. 13.

Paper 8/09657K