DICHIARAZIONE SOSTITUTIVA DI CERTIFICAZIONE E DI ATTO DI NOTORIETA', AU SENSI DEGLI ARTT. 46 E 47 DEL D.P.R. 445/2000, RELATIVA ALLA DETERMINAZIONE ANALITICA DELL'APPORTO INDIVIDUALE DEGLI AUTORI NELLE PUBBLICAZIONI SCIENTIFICHE IN COLLABORAZIONE

Il sottoscritto GUIZZI GUIDO, C.F. GZZGDU75D06F839G, nato a Napoli il 6/4/1975, cittadinanza Italiana, consapevole delle sanzioni penali previste in caso di dichiarazione mendace, ai sensi dell'art. 76 del D.P.R. 445/2000 e s.m.i.

DICHIARA

che la pubblicazione sottoelencata è stata preparata dagli autori operando in stretta collaborazione e con impegno individuale equivalente.

"P. Holimchayachotikul, K. Leksakul, G. Guizzi (2011). Robust Design for Etching Process Parameters of Hard Disk Drive Slider Fabrication Using Data Mining and Multi Response Optimization. WSEAS TRANSACTIONS ON SYSTEMS AND CONTROL, vol. 6, p. 15-24, ISSN: 1991-8763"

Napoli 16 novembre 2012

In Fede.

Robust Design for Etching Process Parameters of Hard Disk Drive Slider Fabrication Using Data Mining and Multi Response Optimization

PONGSAK HOLIMCHAYACHOTIKUL[1]
College of Arts, Media and Technology
Chiang Mai University, Chiang Mai, 50200, THAILAND
Email: holimchayachotikul@hotmail.com[1]

KOMGRIT LEKSAKUL[2]
Department of Industrial Engineering
Faculty of Engineering Chiang Mai University, Chiang Mai 50200, THAILAND
Email: komgrit@eng.cmu.ac.th.[2]

GUIDO GUIZZI[3]
Department of Materials Engineering and Operations Management
University of Naples "Federico II"
P.le Tecchio – 80125 Napoli ITALY
Email guido.guizzi@unina.it[3]

Abstract: - This paper is to provide a proper insight into solving a multi-response optimization problem using hybrid response surface methodology (RSM) based on robust parameter design (RPD) concepts and data mining (DM) for the multi-response optimization of a RIE process. Over the recent years in many high precision manufacturing organizations, domain experience and engineering judgment have been used to handle multiple response optimization problems. These manners have led to the increase in uncertainty during the decision-making process. This situation has also happened in the hard disk drive fabrication based on reactive ion etching (RIE). This process is highly complicated in setting the parameters using the slider to the right customer specifications. Therefore, this paper presents a hybrid model to optimize the concerning responses of this process in terms of mean and variance. The silicon plates with a patterned wet film photo resistance as a base substrate are used to demonstrate this research. To begin the proposed approach, design of experiment (DOE), named central composite design (CCD), was employed to accumulate the process records and to specify the significant parameters of the process. Then, support vector regression (SVR) was brought into play to institute the nonlinear multivariate relationships between process parameters and responses. Data obtained from DOE were used in the training process. Last but not least, the regression decision tree and domain engineering knowledge were opted for the initial point of optimization algorithms as well. In conclusion, the reduced gradient search algorithm, a hill-climbing procedure and desirability function were adapted to the DOE model while grid search and desirability function were adapted to the SVR model to find the optimum parameter setting. SVR was found to be the technique with the highest prominent accuracy performance, so it was selected to construct a RIE process model. Consequently, the optimum condition from the final model has been efficiently enabled to be applied in real production based on its experiment confirmation.

Key-Words: - Reactive Ion Etching (RIE), Response Surface Methodology (RSM), Support Vector Regression (SVR) Multi-response Optimization, Robust Parameter Design (RPD), Data Mining (DM)

1 Introduction

Most of the quality improvements in industrial processes are based on design of experiments (DOE), to optimize and model manufacturing processes. It has been often used to optimize only target of yield factors devoid of concerning the impact from their variance. Therefore, the results of this method might reach the appropriated long term improvement in terms of process stability. To conquer this problem, the robust design concept will be applied to handle it through multi-objective optimization. In addition, the Taguchi method was also used to constructed signal-to-noise ratios: these bring the standard deviation and mean together to give a single response variable to minimize or maximize. It can be more helpful in terms of application because the Taguchi method [4,24] uses the less experiment resource s than the classical DOE; however, some of the statistical analysis through this method encounters the lack of information according to fractional structure in orthogonal array. At this discussion point, the roust parameter design (RPD) in the classical DOE perspective has provided an optimization alternative to the industry that not only give a sounder and more competent logic to design and analysis but also lets us to make use of Taguchi's robust design concept[5]. On the view of statistical experiment design, central composite design (CCD) is a potent technique not only for the prediction of the interested system's responses but also to find the primary optimum process parameter settings to achieve the desired quality of the optimized process [8]. There has been speedy growth in the manufacturing industry driven by the advances of both technologies and computer in the recent decade. This radically changing environment has made the production process much more complicated. Usual analytical methods may not always be suitable with this environment due to the quantity of process variables and the non-linear nature of the problems. To acquire real data for the analysis, authentic experimentations have to be conducted as historical data cannot be used to build a DOE model. This might result in interruption or disruption in the production process.

Currently, problem solving takes advantage of data mining (DM) in terms of robustness, fault tolerance, self-learning, and self-organizing. It has been found to be a good alternative in process modeling as the countless accomplishments such as K-Means clustering and simple additive weight were employed to build the system for machine adjustment decision in hard disk drive arm manufacturing [10]. Data mining modeling was selected to make up of the demand forecasting system for the food supply chain [11]. Many publications have recently been reported. Most of DM has proffered advantages over traditional techniques for the manufacturing field such as the design of fuzzy logic power system stabilizers using genetic algorithms [30]. Lately ANN has been reported in the literature to be inferior to SVM. SVM using structural risk minimization was superior to traditional ANN approach using empirical risk minimization. Besides, the popular feed-forward ANN learning with back-propagation algorithm (BP) by means of gradient descent algorithm has suffered from local minimum problem. To determine center, weights, threshold, and minimizing an upper bound of the expected test error automatically is the preeminent proper of SVM. Hyperplane and the statistical theory were used to in this novel learning machine. Alternatively, some ANN parameters must be determined on a trial-and-error. As a result, optimum solution cannot be guaranteed.

In this paper a novel alternative by integrated SVR, and DOE together was developed to find the optimum setting of production process parameters. A case study of the RIE process parameters for hard disk drive slider for product A was used to demonstrate the proposed approach. The accuracy of the model developed by SVR trained with data collected from DOE and RSM were compared based on the mean absolute percentage error (MAPE) of the overall, training, and testing dataset. This paper is organized as follows. In section 2, the SVR algorithm, central composite designs, robust parameter design (RPD), and literature review are briefly described. In section 3, the problem statement case study of the RIE process parameters is analyzed. In section 4, the methodology of the proposed approach is shown. Results and discussion of the proposed framework are provided in section 5. This section has three main sub-sections. These are the experiment phase, model performance comparison and process optimization. Finally, the conclusion is found in section 6.

2 Background

Support Vector Regression

SVM can be used for either classification or regression tasks. The latter is called support vector regression (SVR) [26,28], which is the type of SVM adopted in this study. The major purpose of SVR is to provide a function $f(x)$ that consists of ε deviation from the actual obtained target y_i for all training sets, { x_l, y_l }, $x \in \Re^n, y \in \Re^n$ with l observations. The initial stage of SVR is described by the linear function using the form $\langle w \bullet x \rangle + b$, then the nonlinear problem is transferred into a linear problem by a nonlinear map $\Phi(x)$ from the low dimensional input space to a higher-dimensional feature space. At the same moment, $f(x)$ is as flat as possible. SVR approximates functions using the following form:

$$f(x) = \sum_{i=1}^{k} \left(\alpha_i - \alpha_i^*\right) \left\langle x_i \bullet x \right\rangle + b \qquad (1)$$

Letting b be the "bias" term and the kernel functions the nonlinear transformation function be $\Phi(\bullet)$ is shown as

$$K(x_i x) = \left\langle \Phi(x_i) \bullet \Phi(x) \right\rangle \qquad (2)$$

Kernel function of the input space is equivalent to the dot product result in high dimensional space in equation (2)

Central Composite Design (CCD)

One of response surface methodology, named central composite designs (CCD) is wildly used in industrial processes. Box and Wilson [2] introduced a method for fitting a second-order response surface [20,21]. CCD is also evolved from its use in sequential experimentation. This involves the use of a two-level factorial or fraction (resolution V) combined with 2^k axial or star points, where k is a number of factors. To sum up, the CCD consists of full factorial points, 2^k star points, and n_c center runs, This CCD model comprises of the liner, quadratic and interaction terms as following:

$$f(x_n) = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n + \beta_{12} x_1 x_2 + ... + \beta_{n-1,n} x_{n-1} x_n + \gamma_1 x_1^2 + ... + \gamma_n x_n^2 \quad (3)$$

CCD requires a considerably less number of experimental runs needed for fitting a second-order model than that required by 2^k full factorial design and Taguchi's orthogonal arrays approach.

Robust Parameter Design Concept

The standard robust parameter design (RPD) was proposed by Genichi Taguchi in the 1980s to handle the RPD problem throughout DOE approach. This approach is mainly focused on the levels of controllable factor selection in process and product to reach two aims: (1) to control the mean of response at the desired level and (2) to minimize the variability around the target.

RPD is one of the main stages in quality systems from Taguchi's perspective. These consist of three stages, namely system design, parameter design and tolerance design. When the set of functional and fiscal factors is invoked, system design is concerned to the make clarity of ideas and product or process prototype design. Parameter design refers to the step when the product of process input parameter settings are chosen that control the system less sensitive to variability from the uncontrollable factors, affecting the system. Thus, the expected goal is reached within this stage, it will result in the system variation having dramatic decrease and the mean also level at the steady changing. Tolerance design is the final stage. It helps to control both mean and variance following the expected result from the previous stages.

Classification and Regression Trees

Decision trees are constructed by algorithms indicated many means of separating the interested datasets into branch-like sub-trees. These sub-trees are the child of a root node at the top of the tree.

Generally the relationship is taken out, then one or more decision rules can be derived that explain the relationships between inputs and response. Rules from the decision tree can be chosen and utilized to show the decision tree, offering a way to visually inspect and express the tree-like diagram of associations that distinguish the input and response values. The standard format of this decision tree modeling way is drawn in Fig. 1. Moreover, decision rules can foresee the values of unseen interested observations, including input values, but might not enclose values for the response values.

C&RT, abbreviated from Classification and Regression Trees, initially portrayed in the academic book [3]. C&RT is a recursive procedure (two subsets are divided, and this continuous mechanism replicates until some other stopping condition is reached). Two subsets are portioned by C&RT means, thus the records within each subset are more much harmonized than in the preceding subset. The same forecaster field may be made of use numerous times at diverse stages in the tree. It applies suitable separating to create the best way of data with missing values.

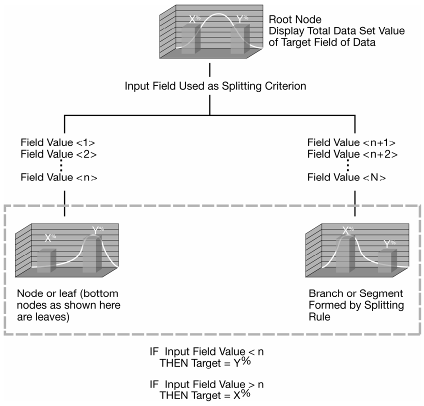


Fig. 1. The example of Decision Tree. [23]

Literature Review

It is certainly true that there is a dramatic increase of the integration between modeling technique based data mining and the design of experiment to industrial process in many aspects. From the literature review in Table 1, most of researchers in this field concentrated on five perspectives to construct the model and optimization for their interested process as follows

1) Solo design of experiment with the optimization algorithms.
   Constructing the model based on actual data from the processes in terms of multi or solo objective functions and then applying the optimization algorithms such as Sequential Quadratic Interpolation, Gradient Descent Method, Conjugate Gradient Method.
2) Fuzzy theory
   To classify and transform the qualitative inputs before constructing the model, analysis of the value of final predictive response is performed before transforming it to the real scale and using this theory to construct the relationship rule between response and input.
3) Data mining for model construction
   To find the suitable component or structure of data mining algorithms or DOE in order to reach the

minimum error of model, neural network component searching by genetic algorithms is one of exples.

4) Genetic algorithm for process optimization
To find the suitable setting of process parameters after data mining or DOE model construction.
5) Construct the model
To use the DOE or historical data for data mining learning processes.

Table 1
The trend of integration between data mining and DOE in term of process optimization since 1990's

Author (Year)	*Process*	*Data*	*Type*	*Method*
Li, T.-S. and Hsu, C.-M. (2010). [19]	Sub-35 nm contact-hole	DOEd	5	NNDOE PSO
Holimchayachotikul and Phanruangron(2009) [12]	Advance Ceramic	DOEd	2	SVM
Holimchayachotikul *et al.* (2009). [13]	Advance Ceramic fabrication	DOEd	2	SVM
Holimchayachotikul *et al.* (2008). [9]	Hard disk drive	DOEd	2	SVM
Ely, G.R., and Seepersad, C.C. (2009). [6]	Process control of welding	AC	2	SVM
Liao, X. *et al.* (2008). [20]	Crash safety	DOEd	4	DOEGA
Lee *et al*. (2007). [25]	design of vehicles Injection molding	DOEd	3	NNGA
Shen *et al*. (2007). [18]	Micro compact heat exchanger	FEMd	4	NNGA
Hou *et al*. (2006). [14]	IC chip-package wire bonding	DOEd	1,4	DOE , NNGA
Huang and Chen. (2006). [15]	Are welding	AC	2	SVMFF
Kim *et al*. (2006). [17]	The iris, the Wisconsin breast cancer, the wine classification	AC	2	NN,NNFF, SVM
Olabi *et al*. (2006). [22]	CO_2 laser welding	DOEd	5	DOENN
Xu *et al*. (2006). [29]	High-technology manufacturing labor productivity	AC	4	SVMGA, NNGA
Chen *et al*. (2005). [4]	Dairy tofu	AC	1	DOEGA
Khaw *et al*. (1995). [16]	The selection of manufacturing operational policies	DOEd	3	NNDOE

Note :
AC: Actual Process data FEMd: Finite Element Method Data
DOE: Design of Experiment DOEd: Design of Experiment Data
NN: Neural Networks SVM: Support Vector Machines
RSM: Response Surface NNFF: Neural Networks with Fuzzy Logic
PSO : Particle swarm optimization
NNDOE: Neural Networks with Design of Experiment
NNGA: Neural Networks with Genetic Algorithm
DOEGA: Design of Experiment with Genetic Algorithm
DOENN: Design of Experiment with Neural Networks
SVMFF: Fuzzy Support Vector Machines with Fuzzy Logic
SVMGA: Support Vector Machines with Genetic Algorithm

3 Problem Statement and Case Study

Currently, the RIE process is still broadly used in the semiconductor industries. RIE is an amalgamation of both physical removal and chemical reactions. The physical removal of the material is frequently achieved by momentum transfer, while the chemical reactions extract the material surface using chemically-reactive gases or plasma. In addition, suitable patterns and etched steps on the air-bearing surface (ABS) of a magnetic head slider can sustain the slider flying on a magnetic disk surface with tremendous low and steady flying height in a hard disk drive. Conventional head sliders are classically made from wafers of a two-phase ceramic, Al-TiC is another name of TiC/Al_2O_3. Conversely, there were the deep research efforts to make use of silicon as the material of the slider substrate for the reason that it can lessen a process step that leads to decrease in processing time. Moreover, it also diminishes the cost by cutting the slider substrate by etching in place of machining to form a slider profile [7]. Moreover, silicon has been found to endow with better characteristics, being a soft material with high thermal conductivity, making it ideal as a slider body which causes less damage to the media during head disk contact [27].

Slider fabrication based on reactive ion etching is a deserved method due to better process control, high etching rate and cleanliness. To achieve high etching quality, it is vital to carefully select etching parameters. Normally, the domain engineers choose the level of the desired etching parameters based on experience or handbook values. Conversely, this does not ensure that the selected etching parameters result in optimal or near optimal etching quality for that particular RIE machine and environment. Thus, several mathematical models have been developed to correlate etching quality with etching parameters.

In the RIE process, there are three necessary parameters: pressure, coil power and platen power. The wall angle and depth is a vital response in the concerning of customer specification in the slider fabrication. In this study, we consider both responses due to the proposed reason.

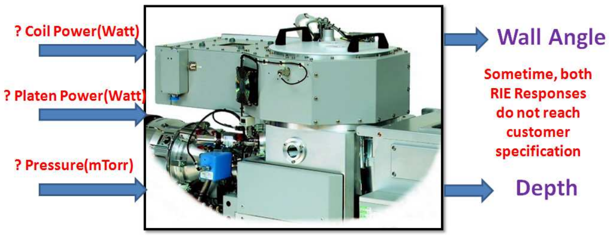


Fig. 2 . RIE Process in company case study.

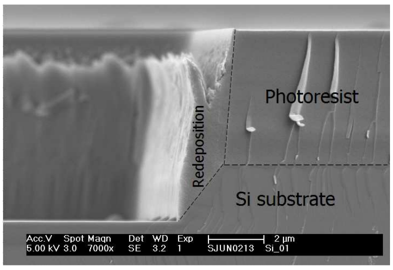


Fig. 3. Example of Si etch profile with redeposition along etch side wall.

All of the etched profiles obtained from this trial experiment are a taper profile feature which is affected by the redeposition phenomenon. The SEM image of the etched profile with redeposition along the etched side wall is shown in Fig. 2 and Fig. 3 . The minimum wall angle in this trial experiment is 47.98 deg and the maximum wall angle is 75.75 deg.

4 Methodology

A schematic diagram of the proposed procedure is shown in Fig. 4. This comprises of the combination of SVR and CCD applied to find the optimum setting of the process parameters in the RIE process for hard disk drive slider for product A. According to the RPD concept to optimize the mean and variance of both responses, both response variances were transformed by the natural log on the grounds that the sample variance does not have a normal distribution (it shown in chi-square scale).

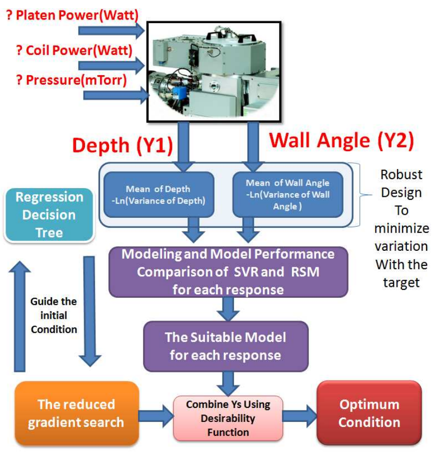


Fig. 4. Schematic diagram of the proposed framework.

Hence, it is typically best analyzing the natural log of variance [20].At this state, we have four responses; two from the mean of responses and the others from variance. The study started from obtaining data from real designed experiments. These data, consisting of input parameters and the corresponding outputs, are then used to train both SVR concurrently (indicated in the computational intelligence learning process). At the same time, a statistical model was developed by CCD using the same data. After the SVR learning process was completed, prediction accuracy of testing data in terms of mean absolute percentage error (MAPE) was used to compare models' performance developed from SVR and CCD.

Regression decision tree, namely CRT, was performed to indicate the initial point for the optimization phase corresponding with domain engineering knowledge. Finally, the grid search method was employed to the best model (in this case was SVR) to find the optimum process parameter setting. The optimum condition from CCD based on the reduced gradient search algorithm and a hill-climbing procedure in the desirability improvement was used to confirm the optimum condition found from grid search.

5 Results and Discussion

5.1 The Experiment Phase

After CCD experiment design with three replications, 60 experiment settings were designed. To perform the experimental design, high, low and axis level of the machining parameters (pressure, coil power and platen power) were selected and shown in Table 2. The experiment was conducted with three replicates and three center points. The machining time for each work piece is constantly 20 minutes. All of the experimental data is shown in Fig. 6. The training data were cases 1-40 and the testing data were cases 41-60.

Table 2
Etching parameters and their levels.

Symbol	Etching parameter	Unit	Level 1	Level 2
X_1	Pressure	mTorr	2	10
X_2	Coil Power	Watt	200	700
X_3	Platen Power	Watt	150	300

After building the final RSM model for each response, the reduced gradient search algorithm and a hill-climbing procedure were employed to maximize the negative natural of log of each response variance and achieve the target optimization of both means of each response. In this study, all computational experiments are performed on an Intel Centrino Core(TM) 2 Duo 2.4 GHz CPU and 3 GB of memory. A total of 60 experimental results obtained from CCD experiment were divided into two sets. The first 80% of data (approximately 48 samples) were used for training, while the rest (20% or approximately 12 samples) were used in the testing process. Accuracy of each model was measured by MAPE. This data set comprises of the input vectors and the corresponding output vector. Firstly, input data was mapped from the input space into a high dimensional feature space using radial basic function kernel (RBF). This function was deservedly used for regression problems [2].The amplitude of RBF was controlled by the vital parameter γ. Unfortunately, the SVR algorithm cannot define the suitable structure by itself; thus, the grid search was applied to find its appropriate component. These results are shown in Table 3. For example, we found that γ of 3.7622 provided the best predictive results and also ε = 0.0031 and c = 579438.2735 from component grid search. These were used in the final SVR model construction for means of wall angle.

Table 3

The suitable SVR structure of each response

Response	γ	c	ε
Mean of Wall Angle	3.7622	579438.2735	0.003126
Mean of Depth	1.073096	220094.9121	0.000031
-Ln(Var Wall Angle)	0.0477	17.7050	0.000023
-Ln(Var Depth)	390.246480	0.0932	0.031623

From both of ANOVA for mean response, we can notice the P-value of square source is significant, so we can ensure that this problem is non-linear. This dataset can be used to construct the model from the non-linear techniques. The example of the estimated regression coefficients from mathematical response model structure and analysis of variance for mean of two responses are as following

Estimated Regression Coefficients for Mean_of_Wall_Angle

Term	Coef	SE Coef	T	P
Constant	125.797	16.0160	7.393	0.000
Pressure	1.331	1.1461	1.161	0.251
Coil Power	0.048	0.0242	1.967	0.055
Platen Power	-0.636	0.1203	-5.289	0.000
Pressure*Pressure	-0.167	0.0387	-4.320	0.000
Coil Power*Coil Power	0.000	0.0000	-10.464	0.000
Platen Power*Platen Power	0.001	0.0002	2.312	0.025
Pressure*Coil Power	-0.006	0.0010	-5.361	0.000
Pressure*Platen Power	0.013	0.0042	3.011	0.004
Coil Power*Platen Power	0.001	0.0001	7.855	0.000

R-Sq(pred) = 80.59% R-Sq(adj) = 85.22%

Analysis of Variance for Mean_of_Wall_Angle

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	361.342	361.342	40.1491	38.79	0.000
Linear	3	119.720	47.286	15.7619	15.23	0.000
Square	3	138.623	138.623	46.2077	44.64	0.000
Interaction	3	102.999	102.999	34.3331	33.17	0.000
Residual Error	50	51.757	51.757	1.0351		
Lack-of-Fit	5	42.482	42.482	8.4965	41.23	0.000
Pure Error	45	9.274	9.274	0.2061		
Total	59	413.099				

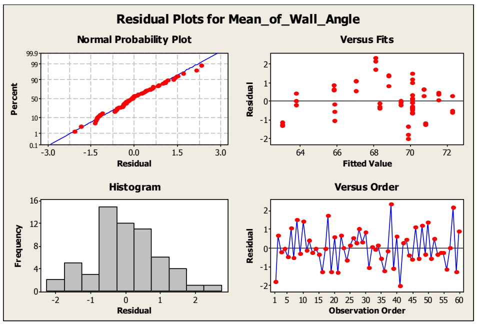
Estimated Regression Coefficients for Mean_of_Depth

Term	Coef	SE Coef	T	P
Constant	-2.47977	1.02834	-2.411	0.020
Pressure	0.30374	0.06926	4.385	0.000
Coil Power	0.00406	0.00146	2.777	0.008
Platen Power	0.01922	0.00727	2.644	0.011
Pressure*Pressure	-0.01723	0.00234	-7.370	0.000
Coil Power*Coil Power	0.00000	0.00000	-2.739	0.009
Platen Power*Platen Power	-0.00003	0.00001	-2.108	0.040
Pressure*Coil Power	0.00002	0.00006	0.249	0.804
Pressure*Platen Power	-0.00065	0.00025	-2.596	0.012
Coil Power*Platen Power	0.00001	0.00001	1.484	0.144

R-Sq = 97.02% R-Sq(pred) = 95.58% R-Sq(adj) = 96.48%

Analysis of Variance for Mean_of_Depth

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	6.15489	6.15489	0.683877	180.89	0.000
Linear	3	5.89463	0.09710	0.032367	8.56	0.000
Square	3	0.22623	0.22623	0.075409	19.95	0.000
Interaction	3	0.03404	0.03404	0.011347	3.00	0.039
Residual Error	50	0.18903	0.18903	0.003781		
Lack-of-Fit	5	0.12860	0.12860	0.025721	19.15	0.000
Pure Error	45	0.06043	0.06043	0.001343		
Total	59	6.34392				



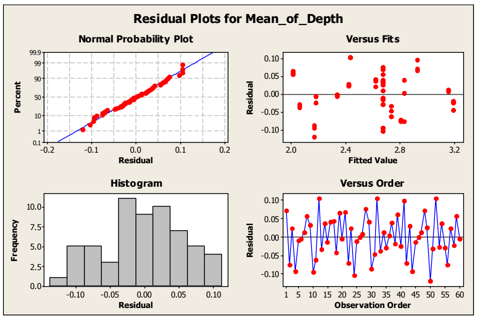


Fig. 5. The graphical residual analysis

From the RSM model we also conduct the residual analysis to prove three main assumptions of DOE as follows: normality assumption, independence assumption and constant variance assumption. The model diagnostic checking can be done easily by graphical analysis of residual as shown in Fig. 5. The tendency of the normal probability plot in each response bends down slightly on the left side and upwards slightly on the right side implies that the tails of error distribution are somewhat thinner than would be anticipated in a normal distribution. Moreover, the residuals of each output are without structure; this plot does not reveal any obvious pattern. So, the assumption of independence is not violated. The plot of residuals versus observation order of each response would not look like an outward-opening funnel or megaphone. So, there is no reason to suspect any violation of the constant variance assumption.

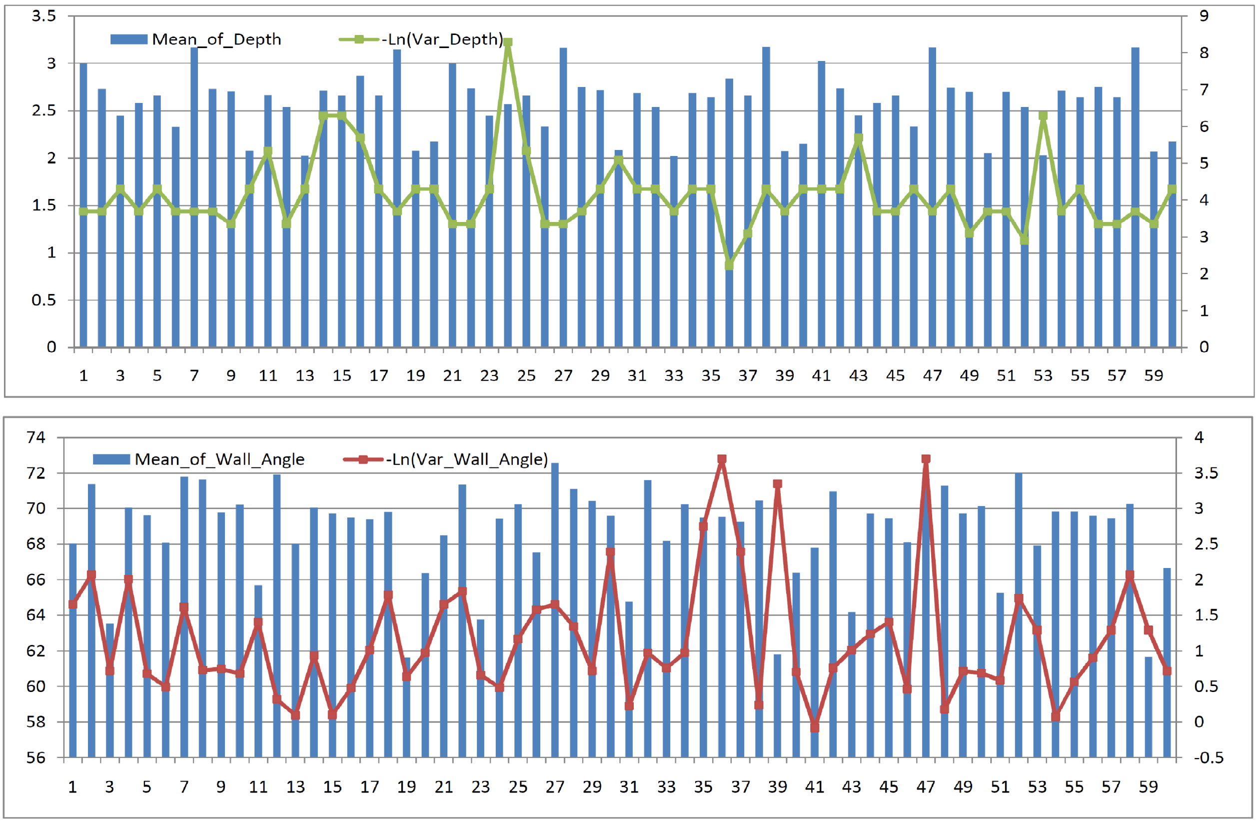


Fig. 6. The four response dataset for CCD and SVR model construction

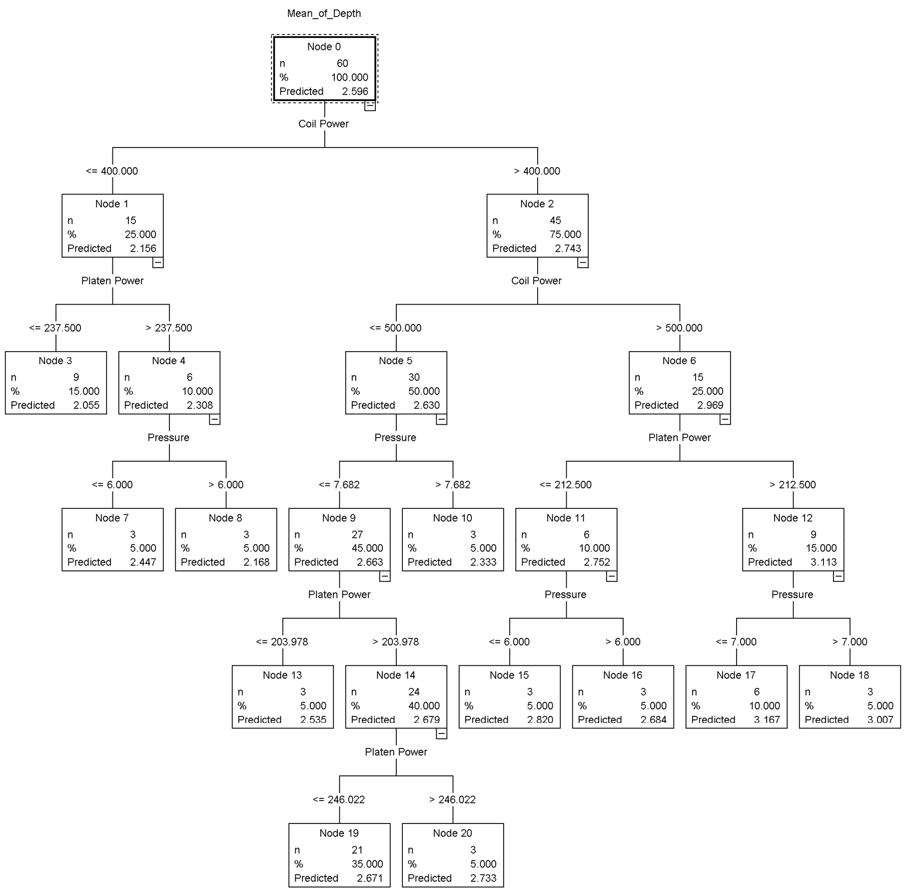


Fig. 7. CRT tree for mean of depth response

5.2 Model Performance Comparison

In this phase, it is vital that to introduce performance evaluation criteria. MAPE in equation 4 was selected as mode performance comparison. Above and beyond, the results from SVR and CCD based on the test set data are shown in Fig. 8. The performance of each model was compared using MAPE of the overall data set, training data set and testing data set. According to the overall data set error, it was found that SVR has a lower error rate than DOE. It is clearly seen that SVR provides good results and is suitable to deal with real problems.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_T}{A_t}\right| \qquad (4)$$

where A_t is the actual value and F_T is the forecast value.

Moreover, the pair t-test of MAPE between the predictive DOE and SVR in test was conducted to determine the model performance difference. From Table 4, we found that the both of the mean response residual from DOE do not differ from that of SVR with 95% of confidence interval. Both of the variance response residuals from DOE are greater than SVR results with 95% of confidence interval. As a result, it was used for further optimization.

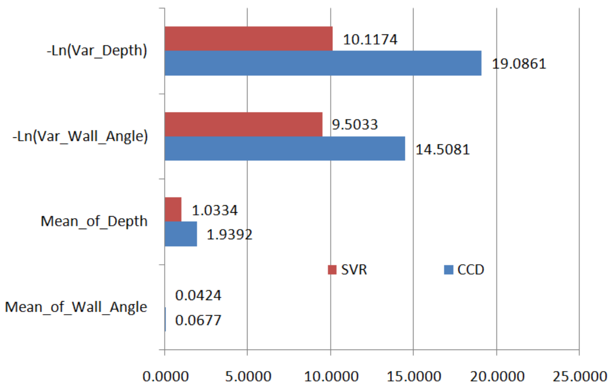


Fig. 8. MAPE comparison between DOE and SVR approach of test set data.

Table 4
The Paired T-Test hypothesis test of residual.

Predictive Response	P-value of Greater than Hypothesis Test
Mean of Wall Angle	0.301
Mean of Depth	0.889
-Ln(Var Wall Angle)	0.000
-Ln(Var Depth)	0.008

5.3 Process Optimization

Before optimization method performing, a CRT tree diagram was drawn to identify the potential range of parameters based on customer specifications (60° of Wall Angle and 2.5 µm of depth). The domain engineers took these four C&RT trees to brainstorm to indicate the initial point for optimization. Four models from all responses were also developed. The example of C&RT tree for mean of depth response is shown in Fig. 7.

From the result of the mean depth response CRT tree, we can simply conclude that the potential range of platen power is more than 237.5 watts and coil power is less than 500 watts. These results lead to the initial solution of optimization. It is 2.5 mTorr of pressure, 250 watts of platen power and 300 watts of coil power.

In this process, SVR, the most accurate model identified in section 5.2, was used to find the optimized parameter settings. The following procedures were adopted. Firstly, composite desirability function (*D*)[5], in equation 5, was assigned weight based on the domain engineer brainstorming for mean of wall angle, mean depth, variance of wall angle and variance of depth as 0.25, 0.25, 0.5 and 0.5 correspondingly.

Secondly, grid search was opted for optimization. Grid search step size for each factor is based on the setting precision. The final optimum condition for the RIE process parameters obtained were from the SVR model at 2.6364 mTorr of pressure, 267.045 watts of platen power and 346.3746 watts of coil power. To simplify optimization results analysis, this condition was put to each SVR response model. These settings result in the prediction for mean of wall angle, mean depth, variance of wall angel and variance of depth following 60.1766, 2.4270, 1.3246 and 4.2020 respectively.

$$D = \left(\prod\left[d_i^{w_i}\right]\right)^{\frac{1}{w}} \qquad (5)$$

where d_i = individual desirability for the ith response,

w_i = weight of the i^{th} response, and $W = \sum w_i$.

$$d_i = f_{\text{target}}(y_i)$$

$$= \begin{cases} 0 & y < L \\ \left(\dfrac{y-L}{T-L}\right) & L < y < T \\ \left(\dfrac{U-y}{U-T}\right) & T < y < U \\ 0 & U > y \end{cases}$$

Note : T = target value L = lower level U = upper level

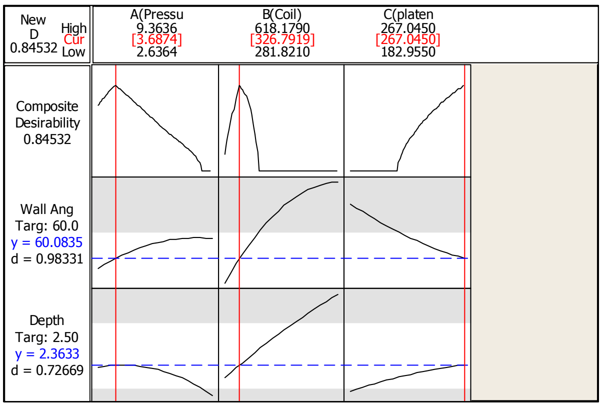


Fig.9. The Optimization Result from CCD

In addition, the variance of this process slightly important response on the ground that these value always in customer requirements. Therefore, we will use only two of mean responses in case of CCD. In the same way, we have found that the optimization of CCD model provides close results following the prediction for mean of depth following 60.08 and 2.36 respectively. Then the desirability function result was drawn in the single graph for domain engineer explanation, thus the example of CCD model results in Fig. 9. The confirmation of SVR condition resulted in closely to the confirmation of CCD conditions of 60.83 and 2.36 for mean of wall angle and mean of depth respectively. Finally. SVR was chosen to model the process in terms of accuracy; however, this condition has to be deployed in the long-run for stability and process capacity analysis in the future before real usage.

6. Conclusion

This paper has described the application of SVR trained with CCD data to obtain a high accuracy modeling of the manufacturing process using RIE process parameters for hard disk drive slider fabrication dataset. Moreover, the modeling of this application uses the output based on the RPD concept. The integration of SVR, DOE, decision tree and grid search to achieve optimization of complex process was proposed. This research has shown the contribution in simple performing and an easy to understand way of optimization in practice such as the initial point for optimization identification using decision trees. It also helps domain engineers to understand more in their process nature. It is certainly true that the performance of SVR depends very much on the data used in the training process. Data with high variation between each trial results in poor modeling performance. Furthermore, the performance depends on the availability of training data. In spite of these limitations, the case study illustrates that SVR can be effectively opted for process modeling and optimization with DoE data. Further potential research may be embraced by combining output vectors from each SVR model using fuzzy desirability functions to achieve optimum conditions or application of other local search methods such as evolution algorithm, particle swarm intelligence, tabu search algorithms. Moreover, we can use the optimum conditions from this proposed study to suggest feasible level determinations of factors in the final process modeling and optimization.

Acknowledgement:

The finance of this research is supported by Department of Materials Engineering and Operations Management of The University of Naples “Federico II” in Italy and The College of Arts Media and Technology, Chiang Mai University, Thailand. The data is also supported by The National Electronics and Computer Technology Center, National Science, Technology Development Agency and Industry/University Cooperative Research Center (I/UCRC) in HDD Components, The Faculty of Engineering, Khon Kaen University and The College of Arts Media and Technology.

References:

[1] D. Anguita, A. Boni, and S. Pace, "Fast training of Support Vector Machines for regression," in Proceedings of the International Joint Conference on Neural Networks, 2000, pp. 210-214.

[2] G. E. P. Box, and Wilson, K. B., , "On the Experimental Attainment of Optimum Conditions, Journal of the Royal Statistical Society," Journal of the Royal Statistical Society, vol. 13, pp. 1-45, 1951 1951.

[3] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression

Trees. New York: Chapman & Hall/CRC, 1984.

[4] M. J. Chen, K. N. Chen, and C. W. Lin, "Optimization on response surface models for the optimal manufacturing conditions of dairy tofu," Journal of Food Engineering, vol. 68, pp. 471-480, 2005.

[5] E. Del Castillo, D. C. Montgomery, and D. R. McCarville, "Modified desirability functions for multiple response optimization," Journal of Quality Technology, vol. 28, pp. 337-345, 1996.

[6] G. R. Ely and C. C. Seepersad, "A comparative study of metamodeling techniques for predictive process control of welding applications," in Proceedings of the ASME International Manufacturing Science and Engineering Conference 2009, MSEC2009, 2009, pp. 395-407.

[7] M. K. Etoh, Tokuyama Mikio and Shiramatsu Toshiya, "Silicon-made magnetic head slider," in United States Patent. vol. 7508631 USA, 03/24/2009.
[8] V. Gunaraj and N. Murugan, "Application of response surface methodology for predicting weld bead quality in submerged arc welding of pipes," Journal of Materials Processing Technology, vol. 88, pp. 266-275, 1999.
[9] P. Holimchayachotikul, Laosiritaworn, W., "Data Ming for Machine Adjustment Decision in Hard Disk Drive Arm Manufacturing : An Empirical Study," in Advances in Soft Computing, Hong Kong, 2009, pp. 459-467.
[10] P. Holimchayachotikul, Phanruangrong, N., "A Framework for Modeling Efficient Demand Forecasting Using Data Mining in
Supply Chain of Food Products Export Industry," in Advances in Soft Computing, Hong Kong, 2009, pp. 1387-1397.
[11] P. Holimchayachotikul, Phanruangrong, N., "Modeling and Optimization of Surface Appearance Defect Reduction for Alumina Substrate Using Design of Experiment and Data Mining Technique," in Advances in Soft Computing, Hong Kong, 2009, pp. 1058-1067.
[12] P. Holimchayachotikul, Phanruangrong, N. , Wanitwattanakosol, J., "Multi-objective optimization of Powdering Process for Alumina Substrate Using DoE and Data Mining Technique," in The 10th Asia Pacific Industrial Engineering & Management Systems Conference (APIEMS 2009) Kitakyushu, Japan: IEMS, 2009.
[13] T. H. Hou, S. H. Chen, T. Y. Lin, and K. M. Huang, "An integrated system for setting the optimal parameters in IC chip-package wire bonding processes," International Journal of Advanced Manufacturing Technology, vol. 30, pp. 247-253, 2006.
[14] X. Huang and S. Chen, "SVM-based fuzzy modeling for the arc welding process," Materials Science and Engineering A, vol. 427, pp. 181-187, 2006.
[15] J. F. C. Khaw, B. S. Lim, and L. E. N. Lim, "Optimal design of neural networks using the Taguchi method," Neurocomputing, vol. 7, pp. 225-245, 1995.
[16] Y. G. Kim, M. S. Jang, K. S. Cho, and G. T. Park, "Performance comparison between backpropagation, neuro-fuzzy network, and SVM," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 3967 LNCS, 2006, pp. 438-446.
[17] H. W. Lee, Y. J. Teng, I. A. Azid, and K. N. Seetharamu, "Neuro-genetic optimization of micro compact heat exchanger," International Journal of Numerical Methods for Heat and Fluid Flow, vol. 17, pp. 20-33, 2007.
[18] T. S. Li and C. M. Hsu, "Parameter optimization of sub-35 nm contact-hole fabrication using particle swarm optimization approach," Expert Systems with Applications, vol. 37, pp. 878-885, 2010.
[19] X. Liao, Q. Li, X. Yang, W. Zhang, and W. Li, "Multiobjective optimization for crash safety design of vehicles using stepwise regression model," Structural and Multidisciplinary Optimization, vol. 35, pp. 561-569, 2008.
[20] D. C. Montgomery, Design and Analysis of Experiments: Wiley, 2008.
[21] R. H. Myer, and Montgomery, D. C., Response Surface Methodology: Process
and Product Optimization Using Designed Experiments. New York: John Wiely & Sons, 1995.
[22] A. G. Olabi, G. Casalino, K. Y. Benyounis, and M. S. J. Hashmi, "An ANN and Taguchi algorithms integrated approach to the optimization of CO2 laser welding," Advances in Engineering Software, vol. 37, pp. 643-648, 2006.
[23] SAS Publishing", "Decision Trees— What Are They?," 2007.
[24] Ross PJ., Taguchi technique for quality engineering. New York: McGraw Hill, 1989.
[25] C. Shen, L. Wang, and Q. Li, "Optimization of injection molding process parameters using combination of artificial neural network and genetic algorithm method," Journal of Materials Processing Technology, vol. 183, pp. 412-418, 2007.
[26] A. J. Smola and B. Schoˆlkopf, "A tutorial on support vector regression," Statistics and Computing, vol. 14, pp. 199-222, 2004.
[27] T. C. R. a. B. C. Stipe, "Method of manufacturing a silicon slider by an alkaline etch," in United States Patent. vol. 7607214 USA, 10/27/2009.
[28] V. N. Vapnik, "An overview of statistical learning theory," IEEE Transactions on Neural Networks, vol. 10, pp. 988-999, 1999.
[29] S. Xu, H. F. Zhao, Z. H. Sun, and X. H. Bao, "Modeling and optimization of high-technology manufacturing productivity," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 3973 LNCS, 2006, pp. 914-920.
[30] Dubey. M, E. Mastorakis. N. "Design of Fuzzy Logic Power System Stabilizers in Multimachine Power System using Genetic Algorithm," WSEAS Transactions on Systems and Control. vol.5 pp 312-322.