IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 12, 2017, accepted April 24, 2017, date of publication April 28, 2017, date of current version June 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2699279

Data-Driven Modeling Using System Integration Scaling Factors and Positioning Performance of an Exposure Machine System

JINN-TSONG TSAI[1], CHENG-CHUNG CHANG[2,3], WEN-PING CHEN[2], AND JYH-HORNG CHOU[2,4,5], (Fellow, IEEE)

[1]Department of Computer Science, National Pingtung University, Pingtung 900, Taiwan
[2]Department of Electrical Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung 807, Taiwan
[3]Metal Industries Research and Development Center, Kaohsiung 811, Taiwan
[4]Institute of Electrical Engineering, National Kaohsiung First University of Science and Technology, Kaohsiung 824, Taiwan
[5]Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung 807, Taiwan

Corresponding author: Jyh-Horng Chou (choujh@kuas.edu.tw)

This work was supported in part by the Ministry of Science and Technology, Taiwan, R.O.C., under Grant MOST 105-2221-E-151-024-MY3, Grant MOST 103-2221-E-153-004-MY2, and Grant MOST 105-2221-E-153-005.

ABSTRACT A data-driven modeling approach is proposed for using system integration scaling factors and positioning performance of an exposure machine system to build models for predicting positioning errors and for analyzing parameter sensitivity. The proposed approach uses a uniform experimental design (UED), multiple regression (MR), back-propagation neural network (BPNN), adaptive neuro-fuzzy inference system (ANFIS), and analysis of variance (ANOVA). The UED reduces the number of experimental runs needed to collect data for modeling. The MR, BPNN, and ANFIS are used to construct positioning models of an exposure machine system. The significant system integration scaling factors are determined by ANOVA. The inputs to the data-driven model are system integration scaling factors f_x, f_y, and f_q, and the output is the positioning error. The UED was used to collect 41 experimental data, which comprised 0.0595% of the full-factorial experimental data. Performance tests demonstrated the excellent performance of the UED in collecting data used to build the MR, BPNN, and ANFIS data-driven models. The data-driven models can accurately predict positioning errors during validation. In addition, a sensitivity analyses of parameters showed that design parameters f_x and f_y have the greatest influence on positioning performance.

INDEX TERMS Data-driven modeling, system integration scaling factors, exposure machine, uniform experimental design, multiple regression, back-propagation neural network, adaptive neuro-fuzzy inference system.

I. INTRODUCTION

Exposure machine systems used for photolithographic processing achieve precision positioning by integrating computer vision, a control strategy, and a servo system. Photolithography is a process used in microfabrication to pattern parts of a thin film of a substrate. It uses light to transfer a geometric pattern from a photomask to a light-sensitive chemical photoresist on the substrate. A series of chemical treatments then either engraves the exposure pattern into, or enables deposition of a new material in the desired pattern upon, the material underneath the photoresist. They also require an automated method of mask alignment for accurate and efficient production. In complex cyber-physical systems, the exposure machine usually performs mask alignment repeatedly until the misalignment meets a predetermined criterion. A large positioning error results in a long processing time. Therefore, an important research issue is how to increase alignment performance by reducing the positioning error. Various methods of increasing alignment performance have been proposed in the literature [1]–[3]. Most researchers have focused on developing automatic alignment algorithms and template matching methods rather than on developing methods for improving system integration parameters. In an auto-alignment system, the purpose of the system integration parameters is to enable applications to be customized so that they can provide different functions in the same machine without a change in the system design. Our previous work integrated an artificial neural network model, a full-factorial experimental design, and a Taguchi-based-genetic algorithm to optimize positional compensation

parameters, i.e., system integration parameters, of an exposure machine [2]. Although experiments showed that the optimized positional compensation parameters decreased the iteration count and alignment time, a full-factorial design requires numerous experiments to collect sufficient data for modeling. The input-output variables for the process have highly complex nonlinear relationships, and estimates of positioning errors require progressively increasing numbers of experiments. Therefore, an effective and efficient method is still needed to build a model of system integration parameters that affect auto-alignment performance in exposure machine systems for predicting positioning errors.

Conventional methods of searching for system integration scaling factors and measuring positioning errors of an exposure machine system include one-factor-at-a-time method and trial-and-error method. However, these methods are not systematic, and they need substantial operator experience and expert knowledge. Another drawback of these methods is that they only obtain feasible system integration scaling factors for an auto-alignment machine. That is, the obtained factors may not optimal. Therefore, this study applied soft computing technology in a systematic data-driven approach to using system integration scaling factors of an auto-alignment system to build models that engineers can use for predicting positioning errors and for analyzing parameter sensitivity.

The proposed approach is performed in three stages. First, a uniform experimental design (UED) [4] is used to gather data for system integration scaling factors and measurement errors. Next, multiple regression (MR) [5], back-propagation neural network (BPNN) [6]–[8], and adaptive neuro-fuzzy inference system (ANFIS) [9]–[11] models are used to construct positioning models of the exposure machine system. Finally, analysis of variance (ANOVA) [12]–[14] is performed to analyze sensitivity and to identify the system integration scaling factors that have the largest effects on positioning precision. In order to build models, for the practical real-world industrial exposure machine system, that engineers can use for predicting positioning errors and for analyzing parameter sensitivity, the novelty of the study is the use of the UED, which has the characteristic of uniform dispersion in a solution space and can reduce the number of experimental runs compared to the full factorial experimental design. Additionally, the well-known approaches, including MR, BPNN, and ANFIS, were used for modeling and mutually verifying the performances of models.

This paper is organized as follows. Section II defines the problem considered in this study. Section III presents the data-driven modeling approach using system integration scaling factors and positioning errors. Section IV presents and discusses the results of an actual engineering design example. Finally, Section V concludes the study.

II. PROBLEM STATEMENT

Figure 1 shows that a vision-based exposure machine is a complex cyber-physical system comprising a vision system, system integration scaling factors, and a servo control

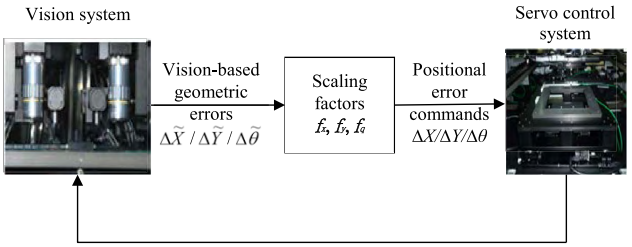

FIGURE 1. Block diagram of vision-based auto-alignment system of an exposure machine system.

system [15], [16]. The vision system includes the cameras, lenses, lighting, and frame grabbers. The servo control system includes the motors, motor drivers, controllers, sensors, and motion control cards. Since the geometric errors ($\Delta\tilde{X}/\Delta\tilde{Y}/\Delta\tilde{\theta}$) obtained by the vision system are not equal to the positional error commands ($\Delta X/\Delta Y/\Delta\theta$) for the servo control system, an exposure machine requires system integration scaling factors, which are a set of parameters f_x, f_y, and f_q between a vision system and a servo feedback control system, to compensate for deviations between $\Delta\tilde{X}/\Delta\tilde{Y}/\Delta\tilde{\theta}$ and $\Delta X/\Delta Y/\Delta\theta$ [15], [16]. The system integration scaling factors enable customized applications because they can be adapted to new functionalities without changing the overall system design. The objective of this study was to develop an effective and efficient method of using system integration scaling factors and positioning performance of an exposure machine system to build models that can be used not only for auto-alignment, but also for predicting positioning errors and analyzing parameter sensitivity.

The auto-alignment system of an exposure machine system performs the following steps to align wafer features with photomask features and to expose the wafer [17].

Step 1: Load a wafer and mask into an exposure machine. Some machines have auto-loading and pre-alignment features while others are very manual in their loading procedures.

Step 2: Align the microscope objectives to the mask alignment marks. Rotate platform mechanism until the microscope objectives are parallel with the alignment marks on the mask, and adjust the X-Y positions on the microscope objectives until they are directly above the alignment marks.

Step 3: Move the wafer alignment mark to the mask alignment mark (Figure 2). For aligners that keep the wafer still and move the photomask over it for alignment, this means that the microscope's position have to be locked in reference to the photomask. When the wafer has finally been aligned completely with the photomask, the wafer should now be moved up into contact with the mask.

Step 4: When the mask and the wafer are correctly aligned, set the exposure time, and expose the wafer. The wafer is exposed for the time set by the

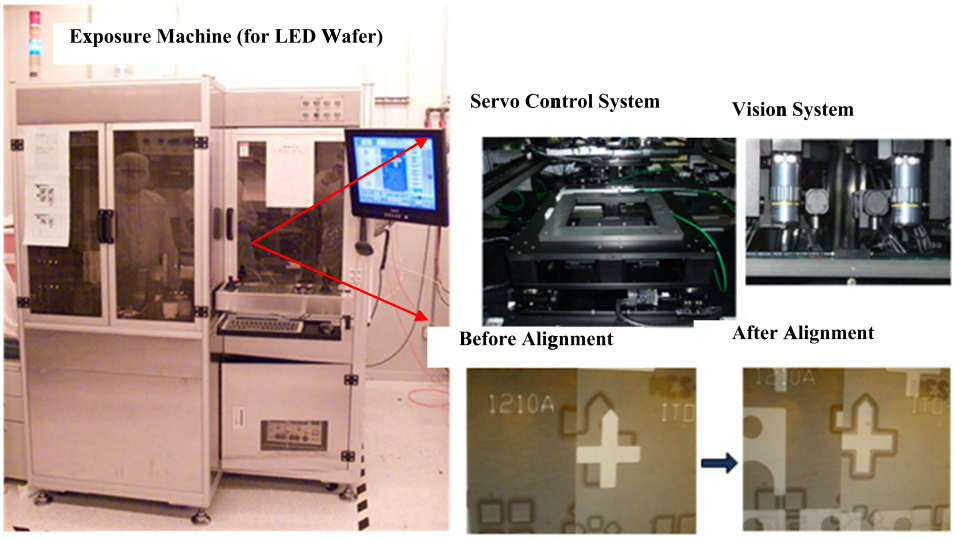

FIGURE 2. Moving a wafer alignment mark to a mask alignment mark.

exposure time. After exposure, the aligners move the wafer away from the mask preparing it for removal from the machine.

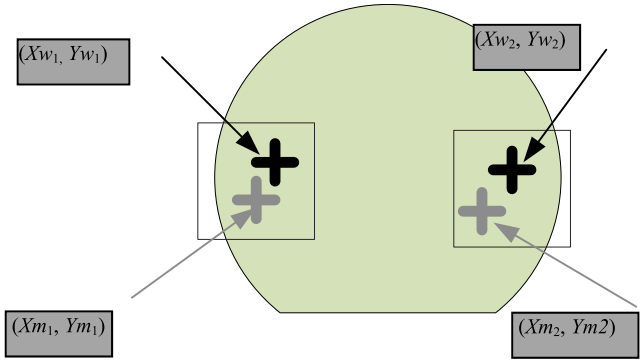

FIGURE 3. The positions between mask alignment marks ((X_{m1}, Y_{m1}) and (X_{m2}, Y_{m2})) and wafer alignment marks ((X_{w1}, Y_{w1}) and (X_{w2}, Y_{w2})).

Figure 3 shows the positions between mask alignment marks ((X_{m1}, Y_{m1}) and (X_{m2}, Y_{m2})) and wafer alignment marks ((X_{w1}, Y_{w1}) and (X_{w2}, Y_{w2})). The $\Delta\tilde{X}$, $\Delta\tilde{Y}$, and $\Delta\tilde{\theta}$ are computed as the displacements of mask alignment marks relative to those of wafer alignment marks. For precise alignment, a set of system integration scaling factors f_x, f_y, and f_q between a vision system and a servo control system is needed to compensate for deviations in ΔX, ΔY, and $\Delta\theta$. In practice, system integration scaling factors are often affected by precision in the assembly of the equipment and platform, precision in machining, and distortion in the lens. The hardware and software in an exposure machine system must be qualified by the engineers, and the proposed data-driven modeling approach can be effective to find the system integration scaling factors. For example, if image resulting from this system contains noisy data either outliers, the measurement errors increase and the performances decline. Additionally, data for system integration scaling factors and positioning performance must be collected for use in building models for predicting positioning errors.

III. DATA-DRIVEN MODELING APPROACH TO BUILDING MODELS FOR PREDICTING POSITIONING ERROR AND ANALYZING PARAMETER SENSITIVITY

The proposed data-driven modeling approach uses system integration scaling factors and positioning performance of an exposure machine system to build models for predicting positioning errors and for analyzing parameter sensitivity. The proposed approach is performed in three stages. First, system integration scaling factors and measurement errors are identified, and UED is used for data collection. Next, MR, BPNN, and ANFIS are used to build positioning models. Finally, ANOVA is applied to analyze the sensitivity of system integration scaling factors. The block diagram and the detailed steps of the proposed data-driven modeling approach are shown in Figure 4 and as follows, respectively.

A. IDENTIFY SYSTEM INTEGRATION SCALING FACTORS AND MEASUREMENT ERRORS, AND USE UED TO COLLECT DATA

A good experimental design should minimize the number of experiments to acquire as much information as possible. The UED developed by Wang and Fang [4] uses space filling designs to construct a set of experimental points uniformly scattered in a continuous design parameter space. The centered L_2-discrepancy (CL_2) is considered by an appealing property that it becomes invariant under reordering the runs, relabeling factors and reflections of the points about any plane passing through the center of the unit cube and parallel to its faces. The latter is equivalent to the invariance of replacing the ith coordinate x_i by 1-x_i for some $i = 1, \ldots, s$.

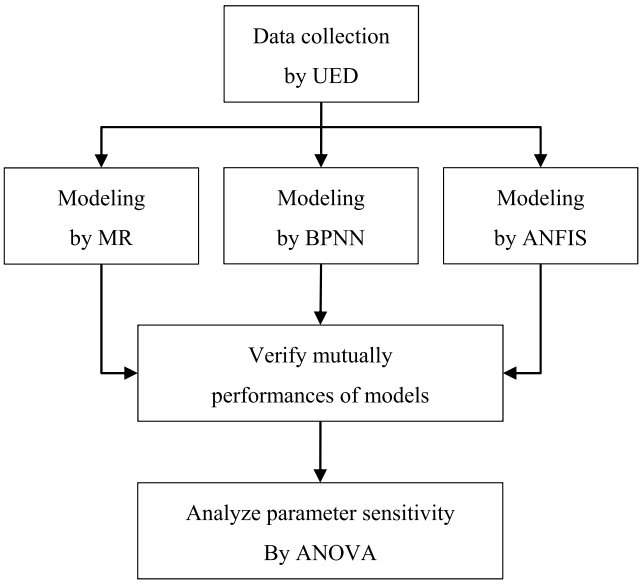

FIGURE 4. The block diagram of the proposed data-driven modeling approach.

For the CL_2, Hickernell [18] has given a mathematically analytical expression as (3.1), as shown at the bottom of this page.

The measure of uniformity of UED has been confirmed by CL_2, and the design points of UED are uniformly scattered in the experimental domain. Therefore, the UED is very suitable for solving problems involving multiple factors with multiple levels and this study used the UED to plan a set of experimental points.

The system integration scaling factors and measurement errors are identified, and the experimental steps are shown below.

Step 1: Identify the main system integration scaling factors.
Step 2: Identify the measurement errors based on the experience of engineers.
Step 3: Achieve an adequate number of uniform design experiments to reflect nonlinear effects, and gather the data.

B. USE MR, BPNN, AND ANFIS TO CREATE SYSTEM MODELS

The statistical MR used for system building in this study included constant, linear, interacting, square, cubic, quartic, and higher terms. The MR uses forward and backward stepwise regression to determine a final model. In each step, the function searches for terms to add to or remove from the model based on the value of the criterion argument. To create a small model, the MR starts with a constant model. To create a large model, the MR starts with a model containing many terms. Compared to a small model, a large model usually has a lower error in terms of fit to the original data. However, a large model may not provide better predictions that are based on new data [5].

This study used a BPNN network that has an input layer, a hidden layer, and an output layer. A BPNN is usually trained in forward and backward phases [6]–[8]. In each run of a BPNN network, training and testing data are used in network training. The number of neurons in the hidden layer is iteratively selected by developing several neural networks and inspecting the mean squared errors of the output. The network training process continues until testing errors increase or until training and testing errors no longer decrease [19], [20]. The BPNN is then designated according to the numbers of inputs, hidden neurons, and outputs. For example, a BPNN 3-8-1 with 3 inputs, 8 hidden neurons, and 1 output is designated BPNN 3-8-1.

This study applied the ANFIS proposed by Jang [9], in which a Sugeno fuzzy model is used for systematically creating fuzzy rules from a given set of input-output data. The structure of this ANFIS design includes a five-layer feedforward neural network. An ANFIS is trained with a hybrid learning algorithm that integrates least squares and gradient descent algorithm. A hybrid procedure comprising a forward pass and a backward pass is performed in each epoch of the ANFIS training algorithm. In the forward pass, a training set is used as input to the ANFIS, and neuron outputs are computed layer by layer. The consequent parameters are determined by the least squares algorithm. During the learning process, the parameters change according to the membership functions. The parameter computation and adjustment are updated by a gradient vector.

The detailed steps are shown below for building a precision positioning model.

Step 1: Use three inputs and one output for building a precision positioning model.
Step 2: Implement uniform design experiments, and create training and testing datasets.
Step 3: Train the MR, BPNN, and ANFIS models and mutually verify the performances of models.
Step 4: Test the performance of the MR, BPNN, and ANFIS models in predicting positioning errors of the exposure machine.

$$CL_2(X) = \left\{ \begin{array}{l} \left(\frac{13}{12}\right)^s - \frac{2}{n}\sum_{k=1}^{n}\prod_{j=1}^{s}\left[1+\frac{1}{2}\left|x_{kj}-\frac{1}{2}\right|-\frac{1}{2}\left|x_{kj}-\frac{1}{2}\right|^2\right] \\ +\frac{1}{n^2}\sum_{k,j=1}^{n}\prod_{i=1}^{s}\left[1+\frac{1}{2}\left|x_{ki}-\frac{1}{2}\right|+\frac{1}{2}\left|x_{ji}-\frac{1}{2}\right|-\frac{1}{2}\left|x_{ki}-x_{ji}\right|\right] \end{array} \right\}^{\frac{1}{2}} \tag{3.1}$$

TABLE 1. U_{41} (41^3) for allocation of design parameters.

Expt. no.	f_x	f_y	f_q
1	1	9	29
2	2	18	17
3	3	27	5
4	4	36	34
5	5	4	22
6	6	13	10
7	7	22	39
8	8	31	27
9	9	40	15
10	10	8	3
11	11	17	32
12	12	26	20
13	13	35	8
14	14	3	37
15	15	12	25
16	16	21	13
17	17	30	1
18	18	39	30
19	19	7	18
20	20	16	6
21	21	25	35
22	22	34	23
23	23	2	11
24	24	11	40
25	25	20	28
26	26	29	16
27	27	38	4
28	28	6	33
29	29	15	21
30	30	24	9
31	31	33	38
32	32	1	26
33	33	10	14
34	34	19	2
35	35	28	31
36	36	37	19
37	37	5	7
38	38	14	36
39	39	23	24
40	40	32	12
41	41	41	41

C. USE ANOVA TO ANALYZE SENSITIVITY OF SYSTEM INTEGRATION SCALING FACTORS

The main objective of the ANOVA [12]–[14] is to determine the variance in results caused by each factor and to determine the total variance caused by all factors. The sum of squares is defined as follows.

$$S_T = \sum_{i=1}^{n} \eta_i^2 - CF, \tag{3.2}$$

where n is the total number of results and $CF = \frac{1}{n}(\sum_{i=1}^{n} \eta_i)^2$.

For factor A, for example, the effect of an individual factor on the variance is calculated as shown below.

$$S_A = \frac{A_1^2}{\text{No.of } A_1} + \frac{A_2^2}{\text{No.of } A_2} + ... + \frac{A_k^2}{\text{No.of } A_k} - CF, \tag{3.3}$$

where A_k is the sum of results and k is the level number.

The variance V_A, F-ratio F_A, and the percentage contribution of factor A are calculated as follows:

$$V_A = \frac{S_A}{f_A}, \tag{3.4}$$

$$F_A = \frac{V_A}{V_{eT}}, \tag{3.5}$$

where f_A is the degrees of freedom for factor A and where V_{eT} is the variance in error.

IV. PRACTICAL INDUSTRIAL EXAMPLE AND IMPLEMENTATION

An actual engineering design example (Figure 2) was used to investigate system integration scaling factors for precision positioning. Based on the training and validation data, system integration scaling factors and positioning performance for an exposure machine system were used to build data-driven models for predicting positioning errors and for analyzing parameter sensitivity.

TABLE 2. Practical experimental data obtained by the U_{41} experiments for training models.

Expt. no.	f_x	f_y	f_q	e_1	θ_1	e_2	θ_2	PEI
1	0.060	0.068	0.168	1005.9885	-103.9545	367.6178	31.346	1.3331
2	0.061	0.077	0.156	987.9866	-111.2902	235.0404	32.7032	1.4682
3	0.062	0.086	0.144	970.7275	-121.0516	138.8727	38.318	1.5404
4	0.063	0.095	0.173	954.3297	-100.1818	34.0839	28.8525	1.6763
5	0.064	0.063	0.161	944.4559	-107.7976	435.9902	19.5518	1.3570
6	0.065	0.072	0.149	927.7727	-116.6478	292.5929	24.5465	1.4742
7	0.066	0.081	0.178	912.8497	-97.8149	152.7114	19.864	1.6296
8	0.067	0.090	0.166	898.9744	-104.9925	69.3756	21.6565	1.7165
9	0.068	0.099	0.154	884.8888	-112.9789	18.8176	23.6346	1.7695
10	0.069	0.067	0.142	877.5123	-122.6496	356.8845	16.727	1.4569
11	0.070	0.076	0.171	863.2241	-101.9458	202.841	12.2016	1.6453
12	0.071	0.085	0.159	848.8851	-110.3649	114.4593	15.4176	1.7255
13	0.072	0.094	0.147	836.6471	-119.62	48.6663	18.6841	1.7856
14	0.073	0.062	0.176	832.498	-100.1049	389.6815	2.3438	1.5085
15	0.074	0.071	0.164	818.7472	-106.6587	257.3855	5.3017	1.6359
16	0.075	0.080	0.152	806.1604	-115.28	161.9262	8.8211	1.7226
17	0.076	0.089	0.140	794.0231	-124.7428	90.0115	12.0191	1.7902
18	0.077	0.098	0.169	782.8767	-103.1619	4.8748	5.6179	1.9393
19	0.078	0.066	0.157	779.2659	-111.273	317.6175	-2.2281	1.5724
20	0.079	0.075	0.145	767.2036	-120.4816	211.1571	1.7282	1.7105
21	0.080	0.084	0.174	756.1915	-100.6004	98.5774	-0.6836	1.8628
22	0.081	0.093	0.162	746.009	-107.8651	36.1648	0.6785	1.9452
23	0.082	0.061	0.150	744.6544	-116.6168	381.0793	-9.8056	1.4041
24	0.083	0.070	0.179	732.914	-97.641	232.6335	-8.9965	1.5905
25	0.084	0.079	0.167	722.203	-104.8634	141.8119	-6.5357	1.7413
26	0.085	0.088	0.155	711.9224	-112.769	74.7595	-5.0624	1.8501
27	0.086	0.097	0.143	702.6273	-122.2593	23.8175	-1.8514	1.9510
28	0.087	0.065	0.172	702.1541	-101.74	287.0932	-15.4446	1.4393
29	0.088	0.074	0.160	690.8695	-109.3918	187.6702	-13.3525	1.6063
30	0.089	0.083	0.148	681.1217	-118.1477	114.0348	-10.6699	1.7423
31	0.090	0.092	0.177	672.1027	-98.6415	32.5866	-11.0609	1.8394
32	0.091	0.06	0.165	674.0674	-105.8691	347.7287	-22.1715	1.2747
33	0.092	0.069	0.153	663.0848	-114.1599	235.7149	-19.1344	1.4769
34	0.093	0.078	0.141	653.3816	-124.0243	155.4968	-16.6606	1.6277
35	0.094	0.087	0.170	645.4565	-103.024	66.1083	-16.3074	1.7393
36	0.095	0.096	0.158	637.8413	-110.6763	19.9789	-15.027	1.8329
37	0.096	0.064	0.146	639.3026	-119.7391	290.927	-25.1196	1.3351
38	0.097	0.073	0.175	629.8792	-100.0412	174.1217	-22.3527	1.5002
39	0.098	0.082	0.163	621.4331	-107.975	103.3282	-20.0823	1.6477
40	0.099	0.091	0.151	612.7824	-116.0543	50.5082	-19.3173	1.7511
41	0.100	0.100	0.180	606.0787	-97.7715	22.1109	-18.1118	1.7783

A. STAGE 1: IDENTIFY SYSTEM INTEGRATION SCALING FACTORS AND MEASUREMENT ERRORS, AND USE UED FOR DATA COLLECTION

The three system integration scaling factors of an exposure machine system are f_x, f_y, and f_q. According to practical engineering experience of the Metal industries Research and Development Centre (MIRDC, http://www.mirdc.org.tw), four measurement errors occur: first position error (e_1) and angle error (θ_1) between the initial and first position points, second position error (e_2) and angle error (θ_2) between the first and second position points. The positioning error index (PEI) is used to integrate $(|e_1| - |e_2|)/|e_1|$ and $(|\theta_1| - |\theta_2|)/|\theta_1|$. The f_x and f_y range from 0.06 to 0.10, and f_q ranges from 0.14 to 0.18. The solution accuracy of each design parameter is 0.001. The number of solutions, i.e., the number of full-factorial combinations of the three design parameters is 68921 ($= 41^3$). To reflect non-linear effects and to collect sufficient data to represent the positioning model, the range of each design parameter is divided into 41 levels. Table 1 shows that this study used U_{41} (41^3) to allocate design parameters. The U_{41} has 41 levels that fit the solution accuracy for f_x, f_y, and f_q. Table 2 shows the practical experimental data obtained by using U_{41} experiments to train system models. To validate the effectiveness of system models, validation data are also generated by uniformly allocating them in a design space of parameters. Table 3 shows a U_{21} (21^3) used to allocate design parameters. Table 4 shows the practical experimental data obtained by using the U_{21} experiments to validate system models.

B. STAGE 2: USE MR, BPNN, AND ANFIS TO CREATE SYSTEM MODELS

The three inputs were f_x, f_y, and f_q. The output was the PEI, which integrates $(|e_1| - |e_2|)/|e_1|$ and $(|\theta_1| - |\theta_2|)/|\theta_1|$. According to the experimental data and the engineering experience of the authors, the system performance improvement obtained by the system integration scaling factors increases as both e_2 and θ_2 approach 0. Therefore, a higher PEI value obtains better performance.

TABLE 3. U_{21} (21^3) for Allocation of Design Parameters.

Expt. no.	f_x	f_y	f_q
1	1	4	5
2	2	8	10
3	3	12	15
4	4	16	20
5	5	20	4
6	6	3	9
7	7	7	14
8	8	11	19
9	9	15	3
10	10	19	8
11	11	2	13
12	12	6	18
13	13	10	2
14	14	14	7
15	15	18	12
16	16	1	17
17	17	5	1
18	18	9	6
19	19	13	11
20	20	17	16
21	21	21	21

Tables 2 and 4 show the datasets used for training and validating the MR model, respectively. The MR model contains constant, linear, interaction, square, cubic, quartic, and higher terms. Model quality is measured by computing the correlation coefficient (R-value) of a linear regression between the target and the model output. An R-value of 1 indicates a perfect correlation. The MR uses stepwise regression to obtain the quartic-order MR model shown below. The quartic-order MR model obtained R-values of 0.99305 and 0.99697 for the training and validation sets, respectively.

$$\begin{aligned} PEI = {} & 17.574 - 1103.6 f_x + 30.637 f_y + 9.2119 f_q \\ & + 198.88 f_x \times f_y - 106.36 f_x \times f_q \\ & + 23315 f_x^2 - 209.68 f_q^2 - 2.0867 \times 10^5 f_x^3 \\ & + 6.7413 \times 10^5 f_x^4 \end{aligned} \tag{4.1}$$

The datasets in Tables 2 and 4 were used for BPNN training and validation, respectively. The momentum was 0.9, and the learning rate was 0.05. The stopping criterion was a training epoch larger than 5000 or a performance goal less than 10^{-6}. The input and output transfer functions were tan-sigmoid and linear functions, respectively. The number of neurons in the hidden layer was iteratively decided by creating several neural networks and inspecting the R-value of a linear regression. Figure 5 shows the three inputs and one output used to train the BPNN architecture. Comparisons of the experimental neural networks showed that the BPNN 3-5-1 network obtained the most accurate positioning results and obtained R-values of 0.99908 and 0.99526 for the training and validation sets, respectively.

The system model was established by the ANFIS using the dataset in Table 2. The initial step size was 0.01, the rate of decrease in step size was 0.9, and rate of increase in step size was 1.1. The stopping criterion was a training epoch larger than 100 or a training error less than 10^{-6}. Both input membership functions were bell-shaped. The output membership function was a linear function. Figure 6 shows the three inputs and one output used to train the ANFIS architecture. The ANFIS model obtained R-values of 1 and 0.98202 for the training and validation sets, respectively.

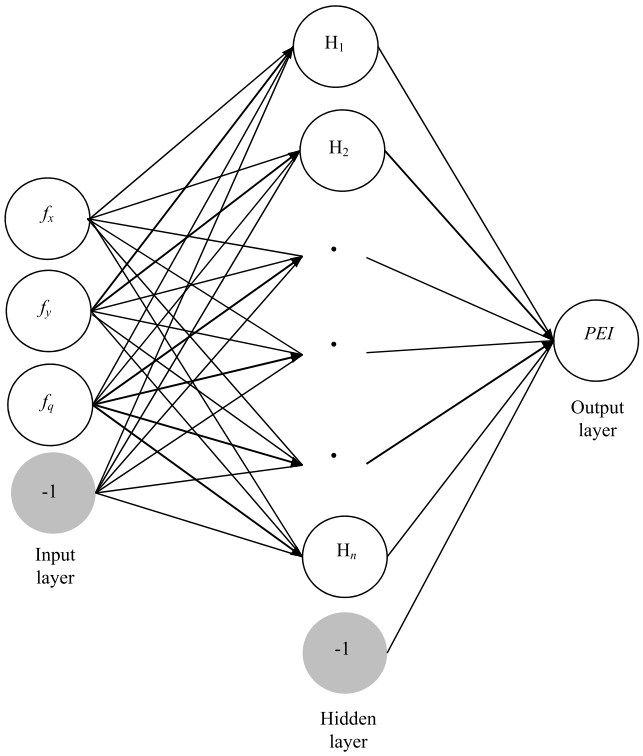

FIGURE 5. Back-propagation neural network with three inputs and one output.

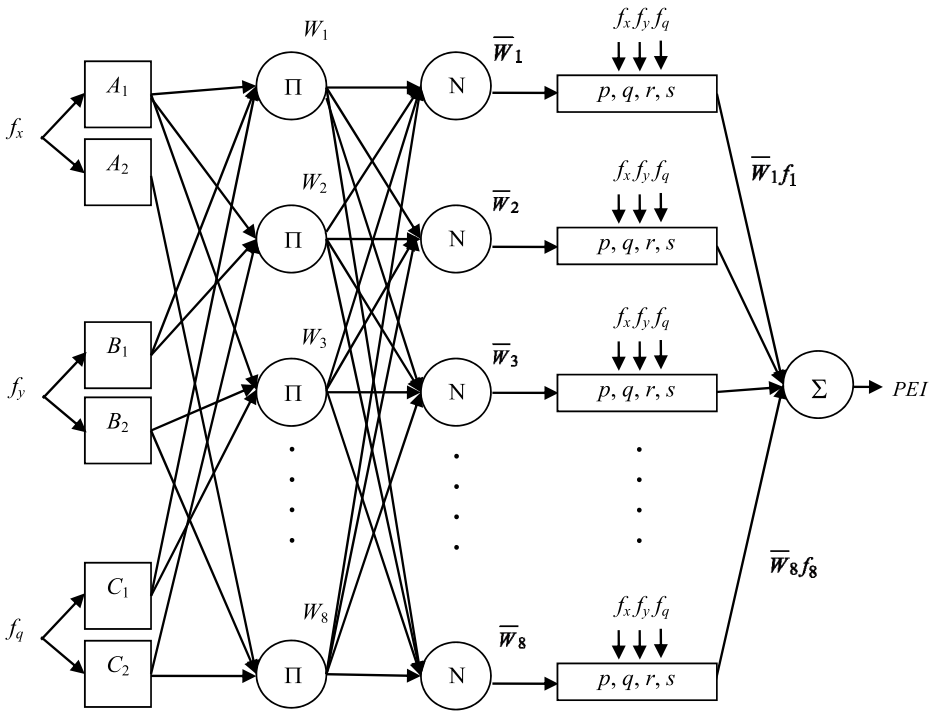

FIGURE 6. Adaptive neuro-fuzzy inference system architecture with three inputs and one output.

The comparisons showed that the quartic-order MR, BPNN 3-5-1, and ANFIS models obtained very similar R-values for the training and validation data. The U_{41} experiments, which were 0.0595% of five-level full-factorial experiments, demonstrated the excellent performance of the model needed to build effective MR, BPNN, and ANFIS models. The prediction results for the U_{21} validation dataset were also very accurate. Figure 7 shows the predication results obtained for the U_{21} validation dataset when the U_{41} dataset was used to build the quartic-order MR, BPNN 3-5-1, and ANFIS data-driven models. Comparisons of positioning error for the U_{21} validation dataset showed that the quartic-order MR, BPNN 3-5-1, and ANFIS models obtained very similar values. Figure 8 further shows that, for the U_{21} validation dataset, the prediction results obtained by the quartic-order MR, BPNN 3-5-1, and ANFIS models were closest to the actual positioning error values. Figure 9 shows that the residuals in the quartic-order MR, BPNN 3-5-1, and ANFIS models were symmetrically distributed near zero with no systematic tendency to appear on the positive or negative sides of the graph between 0.07 to -0.04. The narrower spreads of residual values obtained by the quartic-order MR, BPNN 3-5-1, and ANFIS models were further indications of their superior prediction performance.

C. STAGE 3: USE ANOVA TO ANALYZE SENSITIVITY OF SYSTEM INTEGRATION SCALING FACTORS

Table 5 shows the ANOVA with 95% confidence level in the F-test of three 5-level design parameters obtained using the quartic-order MR model. System integration scaling factors f_x and f_y are statistically significant at the 95% confidence level, which indicates that they are the main cause of variation in positioning performance.

To demonstrate the excellent capability of UED in collecting data for modeling, a five-level full-factorial experimental design was used in performance comparisons. For f_x and f_y, the five levels are 0.06, 0.07, 0.08, 0.09, and 0.10; for f_q, the five levels are 0.14, 0.15, 0.16, 0.17, and 0.18. Table 6 shows the actual experimental data for 125 datasets (F_{125}) obtained by the five-level full-factorial experimental design. The datasets in Tables 6 and 4 were used for model training and validation, respectively.

The MR used stepwise regression to obtain the quartic-order MR model shown below. The quartic-order MR model obtained R-values of 0.99687 and 0.99477 for the training and validation sets, respectively.

$$\begin{aligned} PEI = {} & 37.265 - 2120 f_x + 37.666 f_y + 7.0823 f_q \\ & + 168.35 f_x \times f_y - 65.64 f_x \times f_q - 15.124 f_y \times f_q \\ & + 42671 f_x^2 - 224.95 f_y^2 - 3.7136 \times 10^5 f_x^3 \\ & + 1.1817 \times 10^6 f_x^4 \end{aligned} \tag{4.2}$$

The parameter settings were the same as those shown above for the BPNN and ANFIS. Comparisons of the experimental neural networks showed that the BPNN 3-8-1 network obtained the best performance of the precision positioning system. The BPNN 3-8-1 model obtained R-values of 0.99955 and 0.99814 for the training and validation sets, respectively. The ANFIS model obtained R-values of 0.99921 and 0.99797 for the training and validation sets, respectively.

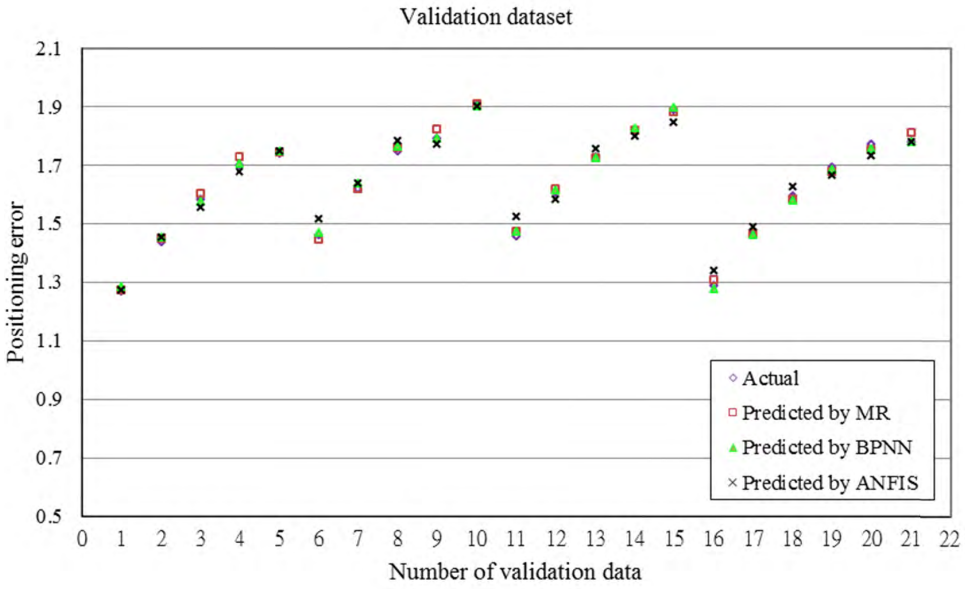

FIGURE 7. Comparison of the actual and predicted values, where the predicted values are obtained by the quartic-order MR, BPNN 3-5-1, and ANFIS data-driven models built using the U_{41} dataset.

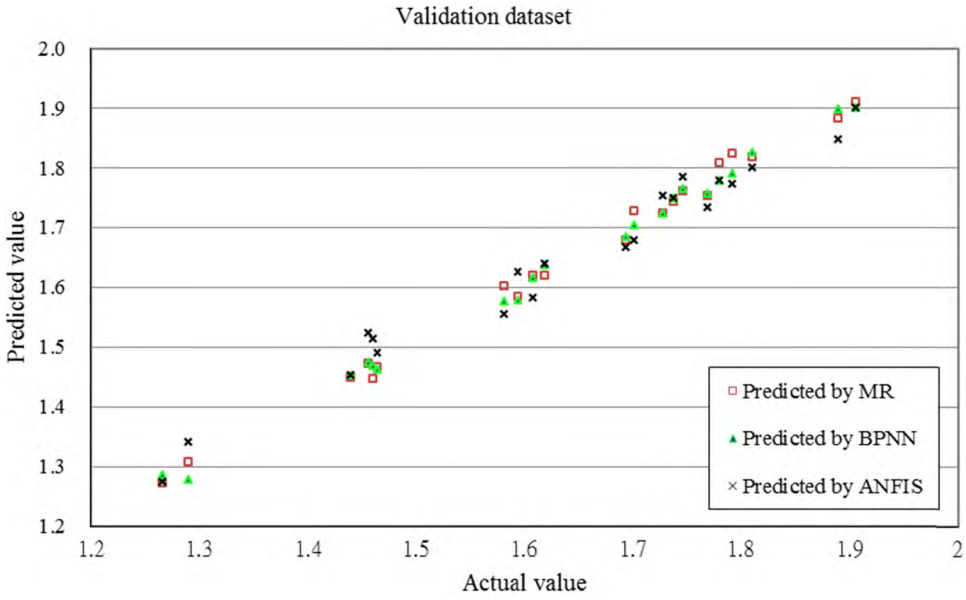

FIGURE 8. Actual values for positioning error compared with the values predicted by the quartic-order MR, BPNN 3-5-1, and ANFIS data-driven models built using the U_{41} dataset.

In Table 7, the RMSEs and R-values for the training (U_{41} and F_{125}) and validation (U_{21}) sets are compared in the MR, BPNN, and ANFIS data-driven models. For each design parameter, the five-level full-factorial experiments covered a solution space accurate to within 0.01. The 125 sets of experimental points in the F_{125} experimental dataset were used to build MR, BPNN, and ANFIS models with high R-values. The models accurately predicted the results of design combinations that appeared in a solution space accurate to within 0.001.

Since the solution accuracy of each design parameter in this study was 0.001, the number of full-factorial combinations for f_x, f_y, and f_q was 68921. In U_{41} (41^3), 41 levels that fit the solution accuracy of f_x, f_y, and f_q were used to construct 41 sets of experimental points, which were uniformly distributed in a solution space accurate to within 0.001. The MR, BPNN, and ANFIS data-driven models that were built using the U_{41} dataset had high R-values and could accurately predict the results for design combinations that appeared in a solution space accurate to within 0.001. The major

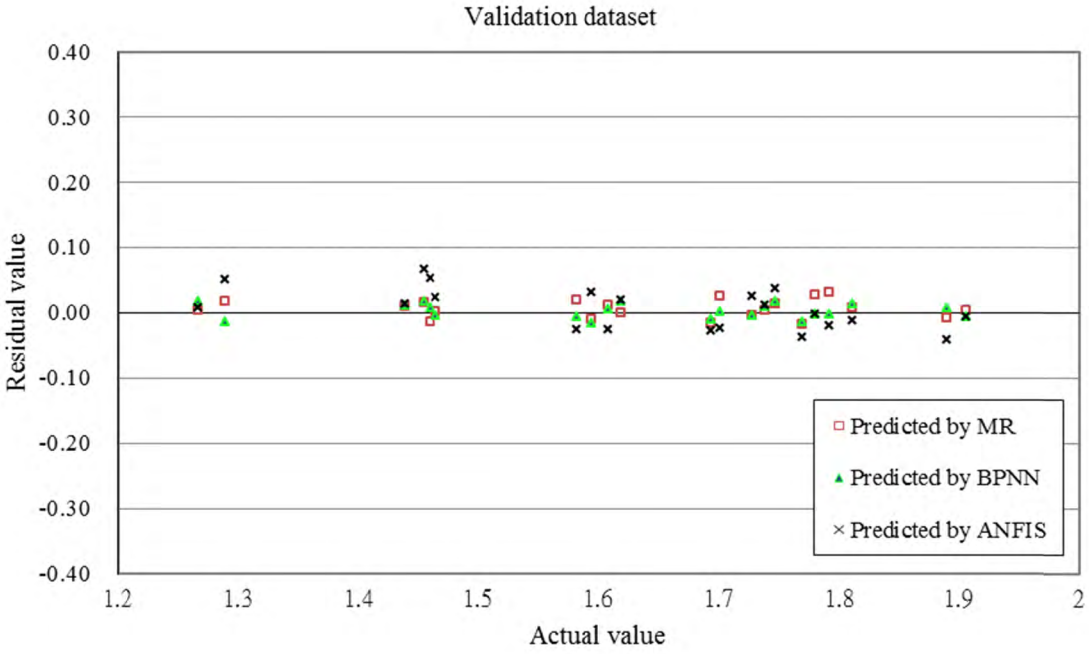

FIGURE 9. Comparison of residual values of positioning errors obtained by the quartic-order MR, BPNN 3-5-1, and ANFIS data-driven models built using the U_{41} dataset.

TABLE 4. Practical experimental data obtained by the U_{21} experiments for validating models.

Expt.no.	f_x	f_y	f_q	e_1	θ_1	e_2	θ_2	PEI
1	0.060	0.066	0.148	1005.6188	-117.9147	423.8306	36.7201	1.2671
2	0.062	0.074	0.158	972.8468	-110.3787	270.2052	31.2273	1.4394
3	0.064	0.082	0.168	941.2954	-104.3613	153.1727	26.6983	1.5815
4	0.066	0.090	0.178	912.1747	-98.2355	65.6174	22.2971	1.7011
5	0.068	0.098	0.146	885.1322	-119.7775	29.3225	27.3303	1.7387
6	0.070	0.064	0.156	866.5379	-112.0082	387.3412	10.3614	1.4605
7	0.072	0.072	0.166	840.2511	-105.1355	251.0133	8.6527	1.6190
8	0.074	0.080	0.176	816.5741	-99.762	145.8245	7.4408	1.7468
9	0.076	0.088	0.144	794.0333	-121.7763	94.9893	10.7208	1.7924
10	0.078	0.096	0.154	773.3438	-113.7483	24.0631	7.1592	1.9060
11	0.080	0.062	0.164	761.945	-106.8286	363.6341	-7.1957	1.4554
12	0.082	0.070	0.174	741.0731	-101.038	238.1355	-7.1041	1.6084
13	0.084	0.078	0.142	721.4959	-123.0072	172.2548	-4.0946	1.7280
14	0.086	0.086	0.152	703.4215	-114.8267	90.6812	-6.9296	1.8108
15	0.088	0.094	0.162	686.7827	-107.7631	28.086	-7.4394	1.8901
16	0.090	0.060	0.172	681.3934	-101.6292	349.501	-20.0345	1.2900
17	0.092	0.068	0.140	663.5808	-124.8405	259.6856	-17.9921	1.4646
18	0.094	0.076	0.150	647.3035	-116.3021	164.6629	-17.5918	1.5943
19	0.096	0.084	0.160	632.2226	-109.4345	90.8796	-17.7473	1.6941
20	0.098	0.092	0.170	618.0942	-102.8407	35.6369	-17.7543	1.7697
21	0.100	0.100	0.180	605.1675	-97.6237	22.0454	-17.8628	1.7806

TABLE 5. Results of ANOVA.

Parameter	*SS*	*DOF*	*V*	*F*	Significance
f_x	8.4589	4	2.11472	32.48	Yes
f_y	31.4369	4	7.85921	120.69	Yes
f_q	0.3688	4	0.0922	1.42	No
Error	0.7814	12	0.06512		
Total	41.0459	24			

SS, sum of squares; *DOF*, degree of freedom; *V*, variance; *F*, *F*-test value; significance is at 95% level confidence.

contribution of this study is the use of the UED that reduced the number of experimental runs compared to the full factorial experimental design method used in Tsai *et al.* [2]. The well-known approaches, including MR, BPNN, and ANFIS, were used for modeling and mutually verifying the performances of models. In this study, although the

TABLE 6. Practical experimental data (F_{125}) obtained by five-level full-factorial experiments.

Expt. No.	f_x	f_y	f_q	e_1	θ_1	e_2	θ_2	PEI
1	0.06	0.06	0.14	1004.0947	-122.7281	552.4869	34.673	1.1673
2	0.06	0.06	0.15	1004.5696	-114.6523	542.0091	31.601	1.1849
3	0.06	0.06	0.16	1005.1109	-107.1129	532.7521	28.4909	1.2040
4	0.06	0.06	0.17	1004.5981	-101.0163	523.3349	25.9779	1.2219
5	0.06	0.06	0.18	1004.6658	-95.4215	515.4478	23.714	1.2384
…	…	…	…	…	…	…	…	
26	0.07	0.06	0.14	865.4033	-123.5692	473.4705	13.003	1.3477
27	0.07	0.06	0.15	866.0526	-115.5715	465.7271	10.2675	1.3734
28	0.07	0.06	0.16	865.569	-108.3687	455.4892	8.6887	1.3936
29	0.07	0.06	0.17	866.075	-101.8942	447.4106	6.9429	1.4153
30	0.07	0.06	0.18	866.1533	-96.1071	441.9142	6.0697	1.4266
…	…	…	…	…	…	…	…	
51	0.08	0.06	0.14	760.7434	-123.2246	413.8875	-6.0176	1.4071
52	0.08	0.06	0.15	760.8742	-115.1033	406.4492	-7.2824	1.4025
53	0.08	0.06	0.16	760.3687	-108.2912	397.263	-7.8195	1.4053
54	0.08	0.06	0.17	760.6148	-101.597	391.705	-8.7539	1.3988
55	0.08	0.06	0.18	761.4209	-96.1905	386.3184	-7.8748	1.4107
…	…	…	…	…	…	…	…	
76	0.09	0.06	0.14	679.6619	-124.2231	368.172	-19.2045	1.3037
77	0.09	0.06	0.15	680.7616	-116.2123	362.6978	-19.8948	1.296
78	0.09	0.06	0.16	680.623	-108.5404	356.5897	-19.9703	1.2921
79	0.09	0.06	0.17	680.767	-102.0467	350.2047	-19.9067	1.2905
80	0.09	0.06	0.18	680.2293	-96.6726	344.2064	-19.0394	1.2971
…	…	…	…	…	…	…	…	
121	0.10	0.10	0.14	605.031	-124.7801	13.5462	-16.3341	1.8467
122	0.10	0.10	0.15	605.0565	-116.3853	12.7126	-16.95	1.8334
123	0.10	0.10	0.16	604.6641	-109.3369	14.7543	-17.4439	1.8161
124	0.10	0.10	0.17	605.1113	-102.7462	17.6627	-17.6574	1.7989
125	0.10	0.10	0.18	604.533	-96.7089	22.5334	-18.6575	1.7698

TABLE 7. The RMSEs and R-values for training and validation sets: comparison of MR, BPNN, and ANFIS data-driven models.

Model	Training set		Validation set	
	RMSE	R-value	RMSE	R-value
	U_{41}		U_{21}	
Quartic-order MR	0.0207	0.99305	0.0152	0.99697
BPNN 3-5-1	0.0076	0.99908	0.0181	0.99526
ANFIS	2.7031×10^{-7}	1	0.0268	0.98202
	F_{125}		U_{21}	
Quartic-order MR	0.0169	0.99687	0.0183	0.99477
BPNN 3-8-1	0.0064	0.99955	0.0122	0.99814
ANFIS	0.0085	0.99921	0.0118	0.99797

MR, BPNN, and ANFIS data-driven models built using F_{125} dataset had slightly superior prediction performance, the number of F_{125} data required for training was triple of the number of U_{41} data required for training. Therefore, the U_{41} dataset has excellent capability in collecting data needed to build effective data-driven models. In industries, due to cost and time limits, they do not allow a lot of experiments, and therefore the UED, which has the characteristic of uniform dispersion in a solution space, can collect a small amount of experiment for modeling. Furthermore, the models built by the MR, BPNN, and ANFIS with the data by the UED have excellent capabilities for predicting.

The main differences between Tsai *et al.* [2] and this study are the use of data collection method. The full factorial experimental design and the UED were used in Tsai *et al.* [2] and the study, respectively. The same modeling approaches in Tsai *et al.* [2] (MLP) and this study (BPNN) were used for modeling. Additionally, the TBGA was used to optimize the positional compensation parameters for the exposure machine in Tsai *et al.* [2], while the ANOVA was applied to analyze parameter sensitivity in the study. This study mainly emphasizes the use of an effective and efficient method for data collection in order to reduce cost and time for the real-world industrial applications. The time complexity of this study compared to Tsai *et al.* [2] is that the number of F_{125} data required for training was triple of the number of U_{41} data required for training.

Remark: The practical results in this study confirmed that the proposed integration approaches can effectively build models for predicting positioning errors and for analyzing parameter sensitivity of an exposure machine system. The obtained data-driven models can accurately predict

positioning errors during validation. Here it should be noticed that the proposed integration approaches have been adopted in the MIRDC (http://www.mirdc.org.tw) and some exposure machine system companies. In addition, the proposed integration methods have immediate real-world applications. This paper is an applications-oriented and interdisciplinary study, and this practical article discusses a new application technique and provides interesting solutions to the exposure machine systems.

V. CONCLUSIONS

The data-driven modeling approach proposed in this study accurately predicts positioning errors by using three system integration scaling factors, f_x, f_y, and f_q. The U_{41} experiments, which are 0.0595% of five-level full-factorial experiments, demonstrate the highly effective in collecting data to build MR, BPNN, and ANFIS models. The MR, BPNN, and ANFIS data-driven models built using the U_{41} experimental dataset, which are uniformly scattered in a solution space accurate to within 0.001, can accurately predict positioning errors for the U_{21} dataset. Sensitivity analyses of the three design parameters showed that system integration scaling factors f_x and f_y had the largest effects on positioning performance. The experimental comparisons in this study also showed that the U_{41} dataset is almost as effective as the five-level full-factorial F_{125} dataset in terms of collecting data for modeling and predicting positioning errors using the U_{21} dataset. The MIRDC (http://www.mirdc.org.tw) has already adopted the proposed integrative and systematic approaches to accurately predict positioning errors of an exposure machine system. That is, the MIRDC has benefited from use of the developed integration methods.

REFERENCES

[1] C. M. Yang, C. C. Wen, S. W. Lin, C. C. Chang, and C. T. Lin, "Application of image servo alignment module design to automatic laminating machine for touch panel," *Smart Sci.*, vol. 1, no. 2, pp. 75–81, Jan. 2016.

[2] J.-T. Tsai, C.-T. Lin, C.-C. Chang, and J.-H. Chou "Optimized positional compensation parameters for exposure machine for flexible printed circuit board," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1366–1377, Dec. 2015.

[3] C. C. Wen and S. W. Lin, "2-phase precision alignment visual feedback control system," *Appl. Mech. Mater.*, vols. 764–765, pp. 587–591, May 2015.

[4] Y. Wang and K. T. Fang, "A note on uniform distribution experimental design," *KexueTongbao*, vol. 26, no. 6, pp. 485–489, 1981.

[5] M. A. Efroymson, "Multiple regression analysis," in *Mathematical Methods for Digital Computers*, A. Ralston and H. S. Wilf, Eds. New York, NY, USA: Wiley, 1960.

[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*, vol. 1, D. E. Rumelhart and J. L. McCleland, Eds. Cambridge, MA, USA: MIT Press, 1986, pp. 318–362.

[7] Y. M. George, H. H. Zayed, M. I. Roushdy, and B. M. Elbagoury, "Remote computer-aided breast cancer detection and diagnosis system based on cytological images," *IEEE Syst. J.*, vol. 8, no. 3, pp. 949–964, Sep. 2014.

[8] C.-M. Lin and E.-A. Boldbaatar, "Fault accommodation control for a biped robot using a recurrent wavelet Elman neural network," *IEEE Syst. J.*, doi: 10.1109/JSYST.2015.2409888.

[9] J. S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 23, no. 3, pp. 665–685, Jun. 1993.

[10] M. J. B. Reddy, D. V. Rajesh, P. Gopakumar, and D. K. Mohanta, "Smart fault location for smart grid operation using RTUs and computational intelligence technique," *IEEE Syst. J.*, vol. 8, no. 4, pp. 1260–1271, Dec. 2014.

[11] J.-T. Tsai, J.-H. Chou, and C.-F. Lin, "Designing micro-structure parameters for backlight modules by using improved adaptive neuro-fuzzy inference system," *IEEE Access*, vol. 3, pp. 2626–2636, 2015.

[12] G. Taguchi, S. Chowdhury, and S. Taguchi, *Robust Engineering*. New York, NY, USA: McGraw-Hill, 2000.

[13] H. M. Hasanien, "Design optimization of PID controller in automatic voltage regulator system using Taguchi combined genetic algorithm method," *IEEE Syst. J.*, vol. 7, no. 4, pp. 825–831, Dec. 2013.

[14] S. F. Hsu, M. H. Weng, R. Y. Yang, C. H. Fang, and J. H. Chou, "An experimental design for processing parameter optimization for cathode arc plasma deposition of ZnO films," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 4, pp. 1588–1593, Oct. 2016.

[15] J. L. Chang, M. W. Hung, and K. C. Huang, "Vision-based multi-axis servo control system," in *Proc. Forum Ind. Appl. AOI Syst.*, 2010, pp. 1–5.

[16] Z. Q. Wen *et al.*, "Development of industrial applications of automatic alignment and assembly technologies," *Electr. Monthly*, vol. 20, pp. 140–148, 2010.

[17] A. R. Hawkins, "Contact photolithographic alignment tutorial," Dept. Elect. Comput. Eng., Brigham Young Univ., Provo, UT, USA, Tech. Rep., 2004.

[18] F. J. Hickernell, "A generalized discrepancy and quadrature error bound," *Math. Comput.*, vol. 67, pp. 299–322, Jun. 1998.

[19] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1999.

[20] I. Sandberg *et al.*, *Nonlinear Dynamical Systems: Feed Forward Neural Network Perspectives*. London, U.K.: Wiley, 2001.

JINN-TSONG TSAI received the B.S. and M.S. degrees in mechanical and electro-mechanical engineering from National Sun Yat-Sen University, Taiwan, in 1986 and 1988, respectively, and the Ph.D. degree in engineering science and technology from the National Kaohsiung First University of Science and Technology, Taiwan, in 2004.

He is currently a Professor with the Department of Computer Science, National Pingtung University, Pingtung, Taiwan. From 1988 to 1990, he was a Lecturer with the Vehicle Engineering Department, Chung Cheng Institute of Technology, Taiwan. From 1990 to 2004, he was a Researcher and the Chief of the Automation Control Section, Metal Industries Research and Development Center, Taiwan. From 2004 to 2006, he was an Assistant Professor with the Medical Information Management Department, Kaohsiung Medical University, Kaohsiung, Taiwan. From 2006 to 2014, he was an Assistant and Associate Professor with the Department of Computer Science, National Pingtung University of Education. His research interests include evolutionary computation, intelligent control and systems, neural networks, and quality engineering.

CHENG-CHUNG CHANG received the B.S. degree in electrical engineering from the National Taiwan University of Science and Technology, Taiwan, in 1992, and the M.S. degree in electrical engineering from the National Yunlin University of Science and Technology, Taiwan, in 1996. He is currently pursuing the Ph.D. degree with the Electrical Engineering Department, National Kaohsiung University of Applied Sciences, Taiwan.

He is currently the Project Manager of the Energy and Agile System Department, Metal Industries Research and Development Center. He hosted over 12 research and development projects, granted five patents, and obtained several awards. In the past ten years, he led a Team to help Taiwan industries to develop over ten kinds of advanced equipment. His research interests include automation, machine vision, image processing, the auto-alignment of high precision machine, and system integration technology.

WEN-PING CHEN received the B.S. degree from the National Taiwan Institute of Technology, Taiwan, in 1992, the M.S. degree from National Sun Yat-Sen University, Taiwan, in 2000, and the Ph.D. degree from the National Kaohsiung University of Applied Sciences in 2008, all in electrical engineering.

He is currently a Professor of the Electrical Engineering Department, National Kaohsiung University of Applied Sciences. His current research interests include the applied of internet of things and the protocol design of networks.

JYH-HORNG CHOU (SM'04–F'15) received the B.S. and M.S. degrees in engineering science from National Cheng Kung University, Tainan, Taiwan, in 1981 and 1983, respectively, and the Ph.D. degree in mechatronic engineering from National Sun Yat-Sen University, Kaohsiung, Taiwan, in 1988. He is currently the Chair Professor of the Electrical Engineering Department, National Kaohsiung First University of Applied Sciences, Taiwan. He has co-authored four books. He has authored over 270 refereed journal papers. He also holds six patents. His research and teaching interests include intelligent systems and control, computational intelligence and methods, automation technology, robust control, and robust optimization. He is also a fellow of the Institution of Engineering and Technology, the Chinese Automatic Control Society, the Chinese Institute of Automation Engineer, and the Chinese Society of Mechanical Engineers. He was a recipient of the 2011 Distinguished Research Award from the National Science Council of Taiwan, the 2012 IEEE Outstanding Technical Achievement Award from the IEEE Tainan Section, the 2014 Distinguished Research Award from the Ministry of Science and Technology of Taiwan, the Research Award and the Excellent Research Award from the National Science Council of Taiwan 12 times, and numerous academic awards/honors from various Societies. Based on the IEEE Computational Intelligence Society (IEEE CIS) evaluation, his ''Industrial Application Success Story'' has received the 2014 Winner of Highest Rank, thus being selected to become the first internationally industrial success story being reported on the IEEE CIS Website.