Taguchi-based Methodology for Determining Process
Model of Injection Molding Using Neural Network

Giheung Choi [1] and G.S. Choi [2]

[1] Department of Mechanical Systems Engineering, Hansung University, 389 Samsun-dong 2-ga
Sungbuk-gu, Seoul, Korea (136-792) TEL: +82-2-760-4322 FAX : +82-2-760-4217

[2] School of Electrical and Electronic Engineering, Seoul City University, 90 Jeonnong-dong,
Dongdaemoon-gu, Seoul, Korea (130-743) TEL: +82-2-210-2526 FAX: +82-2-213-8317

Key Words : Injection Molding, Taguchi Method, Neural Network

Abstract

Implementing CIM (Computer Integrated Manufacturing) often requires models of manufacturing processes to be devised for determining optimal process parameters and designing adaptive control systems. Despite the progress made in analytical (mechanistic) modeling, however, empirical models derived from experimental data are more frequently used in practice. This paper describes the development of a neural network model for injection molding process. The model uses CAE (Computer Aided Engineering) analysis data based on Taguchi method which ensures the effectiveness of the model and the efficient learning by the network. In view of the robust process design, only those input parameters that are not overly sensitive to external disturbances but sensitive enough to injection performance are identified using analysis of variance. The model is compared with the traditional polynomial regression model.

Nomenclature

A_j , j = 1, ..., N_0	CAE analysis outputs
df	**Degree of freedom**
DSI	RMS deviation of SI
DWP	$WP_{ave} - LS_{ave}$
LS	**Linear shrinkage**
$LS_{ave}(WP_{ave})$	Average LS (WP)
F_0	*F* ratio
$N_i(N_0)$	Number of inputs (outputs) nodes in process model
N_n	Total number of finite elements
N_e	Total number of finite element nodes
P_p	Packing pressure
SI	Sink Index
SS	Sum of squares
t_f	Filling time

t_h	Holding time
T_m	Melt temperature
T_c	Coolant temperature
VLS	Variance of LS
WP	Warpage
P_j , j = 1, ..., N_0	Output of the network (or the regression model)
z_i , i = 1, ..., N_i	Input to the network (or the regression model)

Introduction

Recent advancement in digital computing has prompted automation of many conventional manufacturing processes. As a result, a certain degree of maturity in facility level automation (NC, CNC, etc) and data level automation (CAD, CAM, Database) has been accomplished. The direct consequence of such automation is large productivity gains. Further gains in productivity depend on the knowledge level automation of decision making process [1], which requires reliable models to be implemented on the computer.

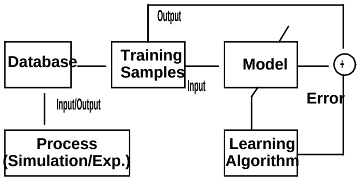
Modeling of manufacturing processes is to establish the mapping relationships between input control parameters and the output performance parameters. Such models can be analytical (mechanistic), empirical or heuristic. Unfortunately, despite the recent progress being made in analytical process modeling, accurate models are not available for many manufacturing processes. Empirical models based on the experimental data are, therefore, most frequently used in practice. One way of developing empirical models is to use statistical regression. In order to develop a regression model, one has to decide each term in the model and collect all the data ahead of time, and then find appropriate coefficients. The model so developed can not be improved incrementally as a set of new data becomes available.

Neural network, however, has many advantages over the conventional statistical techniques. The main advantage of neural network is its self-organizing characteristics which allows the network to adapt to changing environment by incremental learning. Such an adaptive nature enables the on-line adjustment of the model as a set of new data becomes available. It is also suitable for fast computation with parallel architectures, effective in simulating the multi-input/multi-output relation-ship, and has a good generalization capability.

One of the most important areas of automation in manufacturing is the injection molding process because of the associated problems, range of applications, and growing turnover in this field [2]. This requires the complete understanding of the process and the interrelationships among many key factors of the process, which is not realizable at present. In addition, the final properties of injection molded parts are determined in a complex manner by a number of factors, i.e., the geometry of mold cavity, polymeric materials and the process parameters set, to name a few.

The primary geometry of an injection-molded part is usually designed by application engineers to have the necessary functional requirements and the desirable external appearances. Supplementary features such as ribs and bosses are then added to the primary geometry by experienced mold engineers to reinforce the structure and to supplement the functional requirements [3]. Accordingly, the underlying physics of the injection molding process may change such that the chosen process parameters based on the primary geometry deviate from the optimum values for the final geometry. Corrections in process models have to be made to accommodate such changes in process characteristics whenever the supplementary features are added or removed.

This paper describes the development of an empirical model based on neural network for injection molding process. Figure1 depicts a framework of neural network-based injection molding process modeling. Modeling of injection molding process, in general, requires a number of design variables and time-consuming sample data acquisition. In this paper, the experimental data designed by Taguchi method are used for modeling purpose, which ensures the effectiveness of the model and the efficient learning by the network. In view of the robust process design, only those input parameters that are not sensitive to external disturbances but sensitive enough to injection performance are identified using analysis of variance (ANOVA). The present model is then compared with the traditional polynomial regression model to assess its applicability in practice.



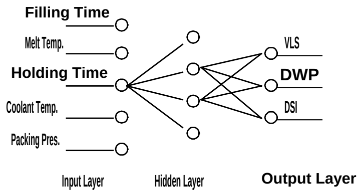
[Figure 1]

Fig. 1 A framework of injection molding process modelling.

In this paper, three functional groups of software are used: A CAE analysis software module developed by the CIMP (Cornell Injection Molding Program) is used for filling, packing, cooling, and warping simulations. A modeling software module, and a user interface and command module which interprets the CAE simulation outputs and interacts with database on demand are also used [4,5].

NEURAL NETWORK-BASED PROCESS MODELING

5-4-3 (or 5-4-1) multi-layered perceptron type back-propagation network with supervised learning algorithm is employed to produce input-output layers as shown in Figure2. The objective of process modeling is to predict the output performance parameters that quantitatively measure the part quality for a variety of given input parameters.



[Figure 2]

Fig. 2 Structure of a 5-4-3 neural network model for injection molding process.

Among many properties that determine the quality of injection molded parts, only the geometrical ones such as the variance of linear shrinkage (hereafter VLS), the difference of warpage (hereafter DWP), and RMS deviation of sink index (hereafter DSI) are used for model outputs. These values are calculated directly from CAE analysis.

DWP is defined as

$$\mathrm{DWP} = WP_{ave} - LS_{ave} \qquad (1)$$

where LS_{ave} and WP_{ave} are the average values of linear shrinkage (LS) and warpage (WP), respectively. Sink index (SI) is an indication of potential shrinkage due to differential cooling and should be minimized, especially near ribs and bosses.

DSI is calculated using the volume as the weighting factor as [6]

$$\text{DSI} = \sqrt{\frac{\sum_i \left(SI_i - SI_{ave}\right)^2}{\sum_i v_i}}\ ,$$

$$i = 1,........, N_n \qquad (2)$$

where v_i and SI_i represent the volume and SI of the i_{th} finite element node, respectively, and SI_{ave} is the average value of SI_i over the total number of nodes in the geometry considered (N_n). DSI indicates the variation of SI around the average value of SI. Typically, a lower DSI is desirable on the part and the runners. Inputs to the network are process parameters such as filling time (t_f), melt temperature (T_m), holding time (t_h), coolant temperature (T_c) and packing pressure (P_p).

Regression Model

Polynomial regression is also performed for all three outputs considered, i.e., VLS, DWP and DSI. If one denotes A_j, j=1,2,3 as a CAE analysis output, it is represented by a regression model as

$$A_j = \phi_j\left(\ t_f\ ,\ T_m\ ,\ t_h\ ,\ T\ \ ,\ P_p\ \right)\ ,$$

$$j = 1, 2, 3 \qquad (3)$$

Model coefficients can be estimated by the standard statistical method.

Design of Experiments

During training of the network, the network accesses a pool of experimental data. These experimental data should be obtained for a broad range of processing parameters to make the applicable range of the network as wide as possible. Using a high dimensional input for the network, however, requires a large number of training samples without necessarily improving the modeling performance. Therefore, in order to facilitate the experimental data collection, the proper design of experiments needs to be devised.

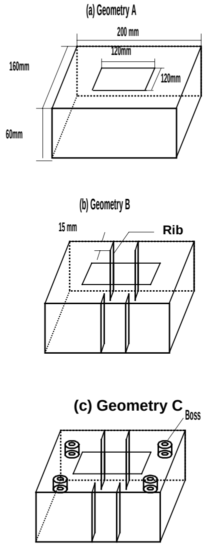
The major goal of designing experiments is to collect as rich a data set as possible within as few experiments as possible. The use of an orthogonal array (fractional factorial design) based on Taguchi method ensures the effectiveness of the model and the efficient learning by the network, keeping the number of experimental runs minimum [7,8]. Furthermore, the robust process design implies the selection of inputs that maximizes the S/N ratio and moves the mean values of output characteristics close to the target values (nominal-the best). For this purpose, analysis of variance (ANOVA) can be performed on the S/N ratios of LS, WP and SI to identify only those input parameters which are not overly sensitive to external disturbances but sensitive enough to injection performance. The S/N ratio of LS, for example, can be defined by

$$\text{S/N}\ Ratio\ = 10\log_{10}\frac{LS_{ave}{}^{2}}{VLS} \qquad (4)$$

ANOVA on the nominal. The best characteristic is to reduce the variability of injection molding performance. Smaller variations in output parameters result in more uniform cooling, less thermally-induced residual stress and less warpage. Consequently, the inputs are optimum when LS, WP and SI have the uniform values over the part being molded.

Simulation Experiments

The Geometry A (primary geometry) considered in this study is shown in Figure3(a). The structure has a dimension of 200x200x60 *mm* with 3 *mm* wall thickness. It also has 120x120 *mm* square hole at the top while the bottom is open space. Geometry B shown in Figure3(b) has 4 ribs where each rib has a dimension of 15x60 *mm*. Geometry C has 4 additional bosses each with a dimension of 10 *mm* inner diameter, 20 *mm* outer diameter, and 20 *mm* height.



[Figure 3]

Fig.3 Geometry considered in this study. 4 ribs are added in Geometry B and 4 additional bosses in Geometry C.

For modeling purpose, all of the inputs to the network were normalized such that [$z_{i,\min}$, $z_{i,\max}$] = [0,1], i=1,.., N_i . A total of 20 LS were calculated from each CAE simulation covering the perimeter of Geometry A and C, and 27 LS for Geometry B. LS_{ave} and VLS were calculated from these data. The warpage was calculated at the midpoints of edges on the top and the bottom faces (4 data each for all geometry). Again, both WP_{ave} and DWP were calculated from these data. The VLS, DWP and DSI were normalized over [0,1] using the predefined maximum and minimum values. A commercially available statistical software package SAS was used for the computation of the regression and the statistical analysis of the constants and the parameters.

Results and Discussion

First, total of 16 training samples, 8 samples each for Geometry A and B, were generated (Data Series A1 and B1, respectively). The operating ranges of these input parameters are given in Table1(1).

	(1)		(2)	
	$z_{\min}$	$z_{\max}$	$z_{\min}$	$z_{\max}$
t_f , sec	1	3	0.5	1.5
$T_{m,}$ ^oC	190	210	210	
t_h , sec	5	10	2.5	
T_c , ^oC	20	40	20	40
P_p , MPa	60	80	60	80

Table1 Operating ranges of input parameters.

These data were based on 2-level (low, high) fractional factorial design in Table2 with 5 input parameters.

No.	1	2	3	4	5	6	7	8
t_f , sec	1	1	1	1	2	2	2	2
$T_{m,}$ ^oC	1	1	2	2	1	1	2	2
t_h , sec	1	2	1	2	1	2	1	2
T_c , ^oC	1	2	1	2	2	1	2	1
P_p , MPa	1	2	2	1	1	2	2	1

Table2 Input parameters based on 2-level orthogonal array. Levels 1 and 2 correspond to z = 0 and 1, respectively.

ANOVA was then performed in a preliminary test. The results of ANOVA shown in Table3 suggest that filling time and packing pressure are the dominant factors for LS, and coolant temperature and packing pressure for WP. ANOVA for SI revealed a similar trend. From this analysis, 3 major factors that affect the injection performance were found to be filling time, coolant temperature, and packing pressure.

	df	(a)			(b)			F (0.05)	F(0.01)
		SS	V	F_0	SS	V	F_0		
t_f	1	33.82	33.82	15.69	182.2	182.2	4.11	7.71	21.2
t_h	1	4.62	4.62	2.14					
T_c	1				333.0	333.0	7.52	7.71	21.2
P_p	1	18.02	18.02	9.35	422.5	422.5	9.54	7.71	21.2
Error	4	8.62	2.61		177.2	44.3			
C Total	7	65.09			1114.8				

Table3 Results of ANOVA for LS, (a), and WP, (b), calculated from Data Series A1 after pooling. Simulation experiments were designed with conditions in Table 2.

Next, a set of experiments was devised using these 3 input parameters, the operating ranges of which are shown in Table1(2). A total of 9 simulation experiments, each one having a combination of different levels of input parameters (low, medium, and high) as shown in Table 4, were carried out for Geometry A (Data Series A2). Similarly, 9 samples each for Geometry B and C were generated based on input parameters listed in Table4 (Data Series B2 and C2, respectively). Melt temperature and holding time were set to the fixed values of 210 ^{o}C and 2.5 sec, respectively.

No.	1	2	3	4	5	6	7	8	9
t_f, sec	1	1	1	2	2	2	3	3	3
T_c, ^{o}C	1	2	3	1	2	3	1	2	3
P_p, MPa	1	2	3	2	3	1	3	1	2

Table4. 9 input parameters based on 3-level orthogonal array. Levels 1, 2, and 3 correspond to z = 0, 0.5, and 1, respectively.

Evaluation samples (6 samples for Data Series EA1, and 7 samples for EA2) were generated from random combinations of input parameters for both Geometry A and Geometry B. The samples were then used to evaluate the performance of models A1, A2, B1, and B2, respectively.

The performance of the network model can be evaluated by the RMS error defined as

$$\text{RMS Error} = \sqrt{\frac{1}{N_o}\sum_{j=1}^{N_o}(A_j - P_j)^2}\ ,\quad j=1,\ldots,N_o \qquad (5)$$

The RMS error by the training samples (Data Series A1) in Figure4 decays very smoothly and saturates after approximately 20,000 iterations for both 5-4-1 and 5-4-3 networks. The steady state error in Figure4 can be viewed as the modeling error of the network.

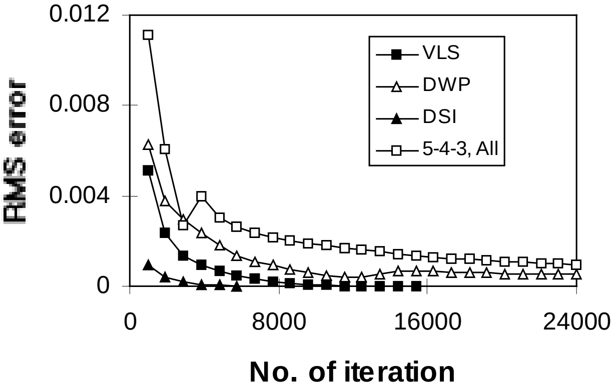
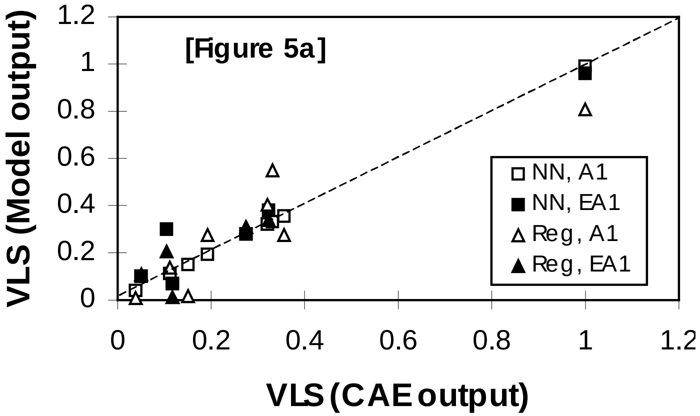
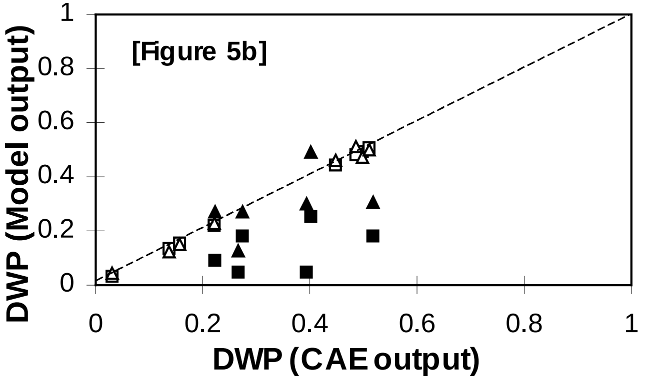


Fig.4 RMS error convergence by the training samples (Data Series A1) for 5-4-1 and 5-4-3 network.

Selected examples of modeling performance with both training samples and evaluation samples are shown in Figure5. The data points for training samples generally fall close to the straight lines, which indicates that the network was well trained to simulate the input/output relationships of the training samples. The overall performance of the neural network models appears to be comparable to that of regression models.





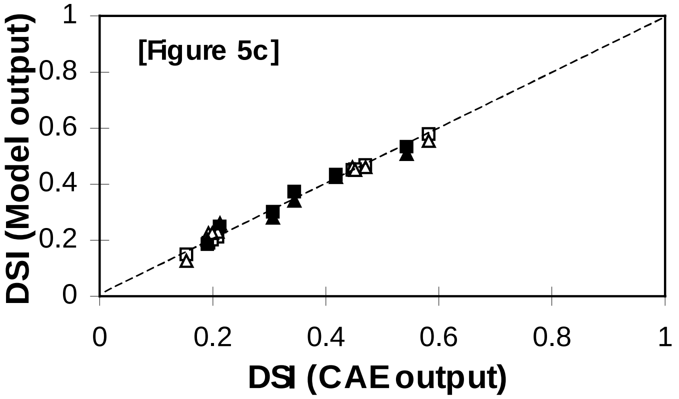


Table5 shows the results of the regression modeling of VLS, DWP and DSI for Data Series A1. R-square value of 0.839 for VLS model indicates that 83.9 % of the variability in VLS, can be explained by the model that uses 5 input parameters as factors. The validity of DWP and DSI regression models are also confirmed by the R-square values larger than 0.9. t value of hypothesis testing indicates a similar trend as predicted by ANOVA.

	Variable	Coefficient t	Standard Error	for Ho	Prob. > | t |
VLS	Constants	-0.3708	0.5396	-0.69	0.563
	t_f	0.3784	0.1591	2.38	0.140
	T_m	-0.1871	0.1591	-1.18	0.361
	t_h	0.2118	0.1591	1.33	0.315
	T_c	-0.1155	0.1591	-0.73	0.543
	P_p	0.1681	0.1591	1.06	0.402
DWP	Constant	0.8535	0.1576	5.41	0.032
	t_f	-0.0300	0.0465	-0.65	0.585
	T_m	-0.0465	0.0465	-1.00	0.423
	t_h	0.0555	0.0465	1.19	0.355
	T_c	-0.3490	0.0465	-7.51	0.017
	P_p	0.0085	0.0465	0.18	0.872
DSI	Constant	-0.0493	0.0753	-0.65	0.58
	t_f	0.2995	0.0222	13.48	0.005
	T_m	0.0366	0.0222	1.65	0.241
	t_h	0.0147	0.0222	0.66	0.576
	T_c	-0.0390	0.0222	-1.76	0.221
	P_p	-0.0535	0.0222	-2.41	0.138

Table5 Coefficient for regression models of VLS, DWP and DSI calculated from Data Series A1.

The generalization capability of the network was tested by presenting a set of inputs from evaluation samples to the network and comparing the outputs from the network with the corresponding outputs from evaluation samples for both VLS and DWP. Modeling error by evaluation samples, however, seems to exceed the acceptable range as illustrated in Figure 5. One way of reducing modeling error in the regression model is to increase the number of terms in the model. However, this is not always possible since the number of training samples is limited when fractional factorial design based on Taguchi method is employed. The regression model could be improved by adding more terms whose forms are yet to be determined. One can also add more hidden nodes in the neural network, which can be done in a compact manner.

Improvement of modeling performance is, however, minimal in most cases [1].

Another way of reducing the modeling error is to increase the number of training samples. Table 6 summarizes the modeling performance of the neural network by both training samples and evaluation samples for Data Series A1.

	8 Samples		16 Samples		32 Samples	
Data Series	A1	EA1	A1	EA1	A1	EA1
VLS	0.00315	0.09209	0.00548	0.19463	0.01547	0.09885
VWP	0.00381	0.23461	0.00393	0.11848	0.02653	0.14190
DSI	0.00307	0.02007	0.00265	0.02231	0.00873	0.03472

Table6 RMS error based on partial factorial design (8 and 16 training samples), and full factorial design (32 samples) in Data Series A1 for 5-4-1 VLS, VWP and DSI models. Data Series EA1 was used for evaluation purpose.

Table7 shows the 16 input parameters based on 2-level orthogonal array. No visible improvement of modeling performance by the full factorial design is observed, which assures the nonlinearity of the process. This prompted the need for training the network with samples based on 3-level orthogonal array.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
t_f, se	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
$T_{m,}$ ^oC	1	1	2	2	1	1	2	2	1	1	1	1	2	2	2	2
t_h, sec	1	2	1	2	1	2	1	2	1	1	2	2	1	1	2	2
T_c, ^oC	1	2	1	2	2	1	2	1	1	2	2	2	1	2	1	2
P_p, MP	1	2	2	1	1	2	2	1	2	1	1	1	2	1	2	1

Table7. 16 input parameters based on 2-level orthogonal array. Levels 1 and 2 correspond to z = 0 and 1, respectively.

For example, Figure6 clearly shows that the modeling error by the evaluation samples is drastically reduced for VLS, DWP and DSI models trained with Data Series A2. R-square values of the regression models for VLS, DWP and DSI in this case were all greater than 0.95.

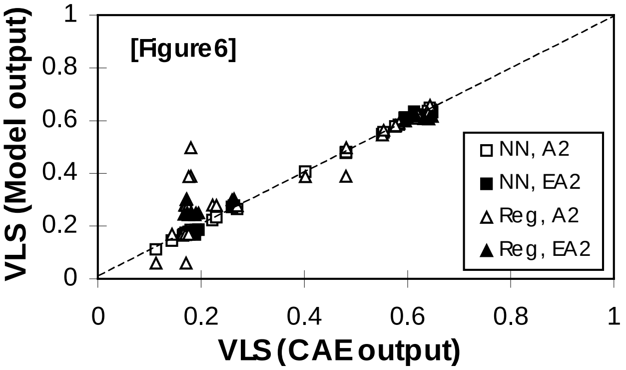
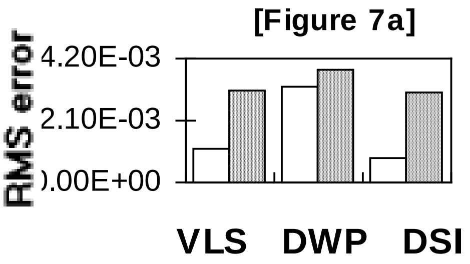


Fig.6 Modeling performance of a 5-4-1 neural network model and a regression model for VLS, DWP and DSI. Both Data Series A2 and EA2 were used for modeling and evaluation, respectively. respectively.

Neural network is particularly effective in simulating the multi-input/multi-output relationship of the process concerned. In our example, adding more output nodes, say from 5-4-1 network to 5-4-3 network, does not substantially affect the overall modeling performance as shown in Figure7.

Therefore, the number of output parameters can be adjusted simply by changing the number of output nodes.



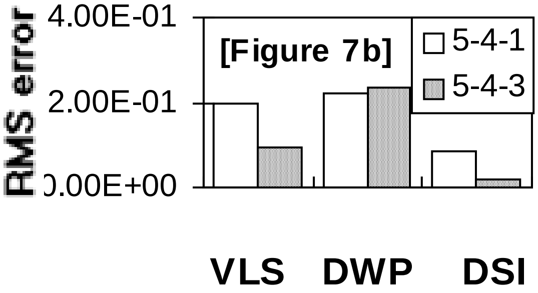


Fig.7 Modeling performance of the neural network for different numbers of out-put nodes. Data Series A1 and A2 were used for (a) and (b), respectively.

Finally, Figure8 shows the variations of VLS and DWP calculated from Data Series A2, B2 and C2. The input conditions are shown in Table 1(2). The values of VLS for Geometry A and B are scattered over the wide range but tend to decrease as ribs are added. No visible variation of DWP is noticed. The values of VLS for Geometry C are, however, widely scattered in the figure.

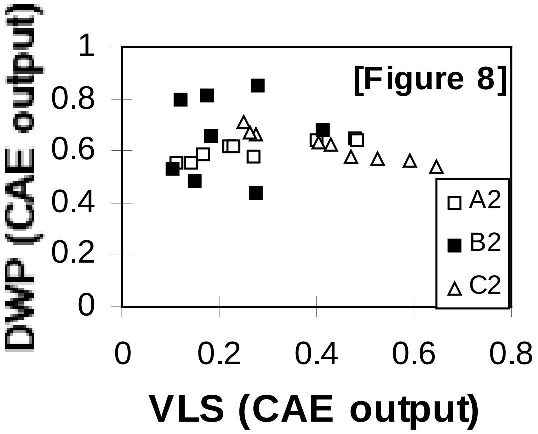


Fig.8 Variations of VLS with DWP calculated from Data Series A2, B2 and C2 for Geometry A, B and C.

This suggests that the corresponding changes in process characteristics are reflected in VLS models. Regression modeling can causes considerable inefficiency in mold design since CAE analysis is required for each change in rib and boss design. The idea of adaptive process modeling based on the self-organizing characteristics of neural network, however, enables the on-line adjustment of the process model with the minimum simulation runs [4].

Summary and Conclusions

In this paper, an effectiveness use of neural network in process modeling of injection molding was demonstrated. The model simulates the relationship between the input process conditions and the output quality measures (VLS, DWP and DSI) and uses the experimental data based on Taguchi method. Taguchi method ensures the effectiveness of process model and the efficient learning by the network, while requiring a minimum number of CAE simulation runs. In addition, in view of the robust process design, only those input parameters that are not overly sensitive to external disturbances but sensitive enough to injection performance were identified using analysis of variance. Analysis of variance must, therefore, be preceded before the process model is actually implemented.

Nonlinearity in the process degrades the efficiency of modeling based on 2-level design. This is manifested by the fact that no visible improvement of modeling performance can be noticed for increasing number of training samples.3-level fractional factorial design, however, improves the modeling performance drastically. The neural network model was then compared with the traditional polynomial regression model to assess the applicability. The neural network model, in general, performs as well as the regression model. However, it has an advantage in dealing with multiple output parameters. Furthermore, in view of implementing a process model at the lower level of expert system the benefit of high processing speed and self-organizing capability of neural network becomes more pronounced as ribs and bosses are added or removed. This allows the network to adapt to changing environment by incremental learning. Such an adaptive nature enables the on-line adjustment of the process model, as a set of new data becomes available.

References

1. Yerramareddy, S., Lu, S., and Arnold, K., 1993, Developing Empirical Models from Observational Data Using Artificial Neural Network. *J. Intelligent Manufacturing*, 4, pp.33-41

2. Actis Dato, M., Di Dio, L., and Godlo, C., 1988, A CIM Concept for Injection Molding. *Kunststoffe German Plastics*, 78(3), pp.9-11

3. Huh Y., and Kim S.G., 1991, A Knowledge-based CAD System for Concurrent Product Design in Injection Molding. *Int. J. Computer Integrated Manufacturing*, 4(4), pp.209-218

4. Choi, G.H., Lee, K.D., Chang N., and Kim S.G., 1996, Adaptive Optimization of Process Parameters of Injection Molding Using Neural Network. *Int. J. Manufacturing Science and Production* (submitted)

5. Choi, G.H., Lee, K.D., Chang N., and Kim S.G., 1994, Optimization of Process Parameters of Injection Molding with Neural Network Application in a Process Simulation Environment. *CIRP Annals*, 43(1), pp.449-452

6. 1993, *C-MOLD User's Guide*, (AC Technology, Inc., NY)

7. Byrne, D. M., and Taguchi, S., 1986, The Taguchi Approach to Parameter Design. *ASQC Quality Congress Transaction*, (Anaheim), pp.168-177

8. Phadke, M.S., 1989, *Quality Engineering Using Robust Design* (Prentice Hall, NJ)