

공정 데이터를 이용한 조업 조건 결정 절차와 사례연구

The Procedure of Finding Operating Conditions Minimizing Quality Loss and Case Study

정일교, 전치혁

포항공과대학교 산업공학과

E-mail: chig@postech.ac.kr

Abstract

The procedure of finding operating conditions minimizing quality loss is proposed with a real industry example. The procedure consists of major two parts – the selection of process variables critical to the response and the determination of operating conditions. The coefficients of ridge regression and the VIP scores of partial least squares are applied to select important process variables. Functional approach and Non-functional approach are used to find proper operating conditions of important process variables.

1. 서론

품질비용을 줄이기 위해서 품질변수와 관련된 중요한 공정인자를 관리하고 조업 조건을 결정하는 문제는 중요하다. 이와 같은 문제를 풀기 위해서 실험계획에 의한 방법이 일반적으로 사용된다. 그러나 실험계획에 의한 방법은 실험비용이 많이 소요되고, 공정인자가 많은 경우에는 실험계획 또한 쉽지가 않다. 요즘은 현장의 조업상태가 체계적으로 수집되고 보관되기 때문에 조업 데이터를 이용하면 보다 경제적으로 중요 공정인자를 찾고 근 최적의 조업조건을 결정할 수 있을 것으로 기대된다. 본 논문은 현장의 조업데이터를 이용하여 중요 공정인자를 선택하고 각 인자의 조업 조건을 결정하는 절차를 실제 사례와 함께 소개하는 것을 목표로 한다.

논문의 구성은 2절에서 분석 공정에 대한 데이터 설명과 분석 절차에 대한 간략한 소개를 한다. 3절에서는 중요 공정인자를 선택하는 방법, 4절에서는 공정인자의 조업 조건을 결정하는 방법을 제시한다. 끝으로 5절에서는 본 연구에 대한 토론 및 추후과제에 대해서 논한다.

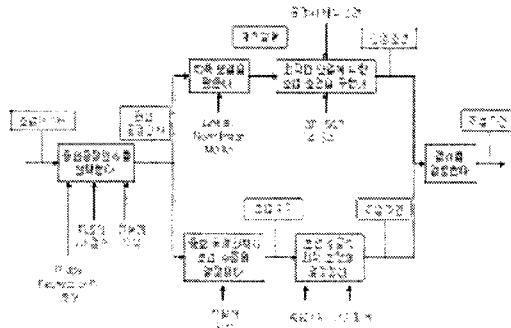
2. 분석 데이터 및 분석 절차

2.1 분석 데이터

분석 대상 공정의 약 1년치 조업데이터를 분석하였다. 현장 데이터는 162개의 공정인자, 3175개의 관측치로 구성되었으며 다수의 결측치가 포함되어 있다. 본 논문에서는 결측치와 공정 설비에 따라서 최종적으로 136개의 공정인자를 분석하며 관측치는 총 1464개이다. 또 품질인자는 당소 특성을 갖는다.

2.2 분석 절차

구체적으로 분석의 목적은 136개의 공정인자 중에서 품질인자에 영향을 미치는 중요 공정인자를 선택하고, 각 중요인자의 조업 조건을 결정하는 것이다. 이를 위해서 본 논문에서는 [그림1]과 같은 분석 절차를 제안한다. [그림1]은 제안된 분석 절차의 Input, Output, 사용기법 및 고려사항 등을 보여준다. 분석 절차는 크게 중요 공정인자 선택과 조업조건 결정방법으로 나눌 수 있는데 3절과 4절에서 자세히 설명한다.



[그림1] 분석절차

3. 중요 공정인자 선택

능형회귀(Ridge Regression)의 계수, PLS(Partial Least Squares)의 VIP(Variable Importance in the Projection) 점수, 그리고 전문가의 지식을 기반으로 중요 공정인자를 선택하고자 한다.

3.1 능형회귀의 계수에 의한 변수 선택

표준화된 데이터에 능형회귀를 적용을 하여 구한 회귀계수의 크기를 변수 선택에 이용한다. 즉, 회귀계수의 절대값 크기가 클수록 품질인자(y)에 영향을 더 많이 미치는 공정인자로 간주한다. 데이터의 표준화는 식(1)을 따른다. 식(1)에서 첨자 i 는 관측치, 첨자 j 는 공정인자에 대한 것이다.

$$\begin{aligned} y_i &\leftarrow (y_i - \bar{y})/s_y, \\ x_{ij} &\leftarrow (x_{ij} - \bar{x}_j)/s_{x_j} \quad (1) \\ (i &= 1, 2, \dots, n; j = 1, 2, \dots, p) \end{aligned}$$

능형회귀 계수는 식(2)를 이용하여 얻는다[2, 3]. 식(2)에서 k 는 편의모수를 나타낸다. k 가 0일 때, 식(2)는 다중회귀의 추정량과 같아진다. 능형회귀는 회귀계수의 추정량에 편의를 가함으로써 불편추정을 회생하는 대신 추정량의 분산을 작게 함으로써 다중 공선성 문제 등을 해결해보고자 하는 것이다.

$$\hat{\beta}_R = (X'X + kI)^{-1} X'y, \text{ where } k \geq 0 \quad (2)$$

편의모수 k 를 결정하는 많은 연구가 진행되고 있는데, 본 연구에서는 Ten-fold Cross-validation에 의해서 k 를 결정한다[2].

3.2 PLS의 VIP 점수에 의한 변수 선택

PLS(Partial Least Squares)는 공정인자들의 선형결합으로 표현 되는 새로운 변수들로 품질인자를 설명한다. 여기서 새로운 변수를 잠재변수라 명명하자. PLS 모델은 식(3)과 같이 표현되며, 여기서 잠재변수 t 는 X 들의 벡타의 선형조합으로 산출된다. 이 때 벡타 x_i 에 대응하는 가중치로 벡타 x_i 를 벡타 y 에 투영시킬 때 얻는 길이를 이용한다. 이는 NIPALS 알고리즘으로 구할 수 있는데 [1]을 참고한다. 또 잠재변수의 개수 (A)는 Ten-fold Cross-validation에 의해서 결정한다.

$$\begin{aligned} X &= t_1 p_1^t + t_2 p_2^t + \dots + t_A p_A^t + E \\ y &= t_1 q_1 + t_2 q_2 + \dots + t_A q_A + f \end{aligned} \quad (3)$$

PLS 모델에서 공정인자 x_j 의 모델에 대한 기여도를 평가하기 위해서 VIP를 식(4)와 같이 정의하며, 보통 VIP가 1보다 큰 공정인자를 의미 있는 변수로 간주한다. 식(4)에서 p 는 공정인자의 개수이다.

$$VIP_j = k \sum_{i=1}^A SS(q_i t_i) w_{ij}^2 / \sum_{i=1}^A SS(q_i t_i), \quad (4) \\ (j = 1, 2, \dots, p)$$

3.3 중요 공정인자 선택 결과

능형회귀의 결과, VIP점수, 및 전문가 지식을 기반으로 중요 공정인자를 다음과 같이 선택한다.

중요 공정인자 = { 능형회귀에 의해서 선택된 중요 공정인자 상위 30개} \cap {VIP 점수에 의해서 선택된 중요 공정인자 상위 30개} + 전문가 지식

Ten-fold Cross-validation에 의해서 능형회귀의 경우, 편의모수를 0.27로 결정했고, PLS의 경우, 잠재변수를 6개 사용하였다. [표1]은 능형회귀 계수의 크기와 VIP 점수의 크기에 따라 나열한 공정인자 상위 30개에 대한 각각의 결과이다. 두 방법에 의해서 x7, x15, x16, x52, x41, x86, x95, x142가 뽑혔으며 전문가의 지식이 반영되어 x13이 추가되고 x41과 x86이 제외되었다. 최종적으로 136개 공정인자 중에서 7개(x7, x13, x15, x16, x52, x95, x142)가 중요 공정인자로 선택되었다.

[표1] 능형회귀 계수의 크기 및 VIP 점수에 의해서 선택된 중요 공정인자 상위 30개

공정인자	회귀계수	VIP	중요 공정인자
x1	0.1234	0.1234	○
x2	0.1234	0.1234	○
x3	0.1234	0.1234	○
x4	0.1234	0.1234	○
x5	0.1234	0.1234	○
x6	0.1234	0.1234	○
x7	0.1234	0.1234	○
x8	0.1234	0.1234	○
x9	0.1234	0.1234	○
x10	0.1234	0.1234	○
x11	0.1234	0.1234	○
x12	0.1234	0.1234	○
x13	0.1234	0.1234	○
x14	0.1234	0.1234	○
x15	0.1234	0.1234	○
x16	0.1234	0.1234	○
x17	0.1234	0.1234	○
x18	0.1234	0.1234	○
x19	0.1234	0.1234	○
x20	0.1234	0.1234	○
x21	0.1234	0.1234	○
x22	0.1234	0.1234	○
x23	0.1234	0.1234	○
x24	0.1234	0.1234	○
x25	0.1234	0.1234	○
x26	0.1234	0.1234	○
x27	0.1234	0.1234	○
x28	0.1234	0.1234	○
x29	0.1234	0.1234	○
x30	0.1234	0.1234	○
x31	0.1234	0.1234	○
x32	0.1234	0.1234	○
x33	0.1234	0.1234	○
x34	0.1234	0.1234	○
x35	0.1234	0.1234	○
x36	0.1234	0.1234	○
x37	0.1234	0.1234	○
x38	0.1234	0.1234	○
x39	0.1234	0.1234	○
x40	0.1234	0.1234	○
x41	0.1234	0.1234	○
x42	0.1234	0.1234	○
x43	0.1234	0.1234	○
x44	0.1234	0.1234	○
x45	0.1234	0.1234	○
x46	0.1234	0.1234	○
x47	0.1234	0.1234	○
x48	0.1234	0.1234	○
x49	0.1234	0.1234	○
x50	0.1234	0.1234	○
x51	0.1234	0.1234	○
x52	0.1234	0.1234	○
x53	0.1234	0.1234	○
x54	0.1234	0.1234	○
x55	0.1234	0.1234	○
x56	0.1234	0.1234	○
x57	0.1234	0.1234	○
x58	0.1234	0.1234	○
x59	0.1234	0.1234	○
x60	0.1234	0.1234	○
x61	0.1234	0.1234	○
x62	0.1234	0.1234	○
x63	0.1234	0.1234	○
x64	0.1234	0.1234	○
x65	0.1234	0.1234	○
x66	0.1234	0.1234	○
x67	0.1234	0.1234	○
x68	0.1234	0.1234	○
x69	0.1234	0.1234	○
x70	0.1234	0.1234	○
x71	0.1234	0.1234	○
x72	0.1234	0.1234	○
x73	0.1234	0.1234	○
x74	0.1234	0.1234	○
x75	0.1234	0.1234	○
x76	0.1234	0.1234	○
x77	0.1234	0.1234	○
x78	0.1234	0.1234	○
x79	0.1234	0.1234	○
x80	0.1234	0.1234	○
x81	0.1234	0.1234	○
x82	0.1234	0.1234	○
x83	0.1234	0.1234	○
x84	0.1234	0.1234	○
x85	0.1234	0.1234	○
x86	0.1234	0.1234	○
x87	0.1234	0.1234	○
x88	0.1234	0.1234	○
x89	0.1234	0.1234	○
x90	0.1234	0.1234	○
x91	0.1234	0.1234	○
x92	0.1234	0.1234	○
x93	0.1234	0.1234	○
x94	0.1234	0.1234	○
x95	0.1234	0.1234	○
x96	0.1234	0.1234	○
x97	0.1234	0.1234	○
x98	0.1234	0.1234	○
x99	0.1234	0.1234	○
x100	0.1234	0.1234	○
x101	0.1234	0.1234	○
x102	0.1234	0.1234	○
x103	0.1234	0.1234	○
x104	0.1234	0.1234	○
x105	0.1234	0.1234	○
x106	0.1234	0.1234	○
x107	0.1234	0.1234	○
x108	0.1234	0.1234	○
x109	0.1234	0.1234	○
x110	0.1234	0.1234	○
x111	0.1234	0.1234	○
x112	0.1234	0.1234	○
x113	0.1234	0.1234	○
x114	0.1234	0.1234	○
x115	0.1234	0.1234	○
x116	0.1234	0.1234	○
x117	0.1234	0.1234	○
x118	0.1234	0.1234	○
x119	0.1234	0.1234	○
x120	0.1234	0.1234	○
x121	0.1234	0.1234	○
x122	0.1234	0.1234	○
x123	0.1234	0.1234	○
x124	0.1234	0.1234	○
x125	0.1234	0.1234	○
x126	0.1234	0.1234	○
x127	0.1234	0.1234	○
x128	0.1234	0.1234	○
x129	0.1234	0.1234	○
x130	0.1234	0.1234	○
x131	0.1234	0.1234	○
x132	0.1234	0.1234	○
x133	0.1234	0.1234	○
x134	0.1234	0.1234	○
x135	0.1234	0.1234	○
x136	0.1234	0.1234	○
x137	0.1234	0.1234	○
x138	0.1234	0.1234	○
x139	0.1234	0.1234	○
x140	0.1234	0.1234	○
x141	0.1234	0.1234	○
x142	0.1234	0.1234	○
x143	0.1234	0.1234	○
x144	0.1234	0.1234	○
x145	0.1234	0.1234	○
x146	0.1234	0.1234	○
x147	0.1234	0.1234	○
x148	0.1234	0.1234	○
x149	0.1234	0.1234	○
x150	0.1234	0.1234	○
x151	0.1234	0.1234	○
x152	0.1234	0.1234	○
x153	0.1234	0.1234	○
x154	0.1234	0.1234	○
x155	0.1234	0.1234	○
x156	0.1234	0.1234	○
x157	0.1234	0.1234	○
x158	0.1234	0.1234	○
x159	0.1234	0.1234	○
x160	0.1234	0.1234	○

4. 공정인자의 조업조건 결정

4.1 함수적 방법에 의한 조업조건

4.1.1 예측모델

7개의 중요 공정인자의 교호작용을 모두 고려하여 단계별 회귀분석에 의해서 식(5)와 같은 예측모델을 만들었다. [그림2]는 예측모델에 대한 관련 통계량을 보여준다.

$$y = -10.6 - 0.233x_{13} - 3.12x_{16} + 0.0133x_{95} + 0.842x_7x_{52} + 0.0861x_7x_{142} + 0.170x_{15}x_{52} \quad (5)$$

Predictor	Coef	TS Coef	T	P	VIF
Constant	-10.624	1.741	-19.24	0.000	
x ₁₃	0.0133	0.001765	11.43	0.000	1.1
x ₁₆	-3.1203	-0.02563	-6.455	0.000	3.2
x ₉₅	0.0133	0.001765	6.19	0.001	1.9
x ₁₅ x ₅₂	0.0133	0.001765	4.69	0.000	1.2
x ₇	-3.1203	-0.02563	-2.21	0.267	1.1
x ₁₄₂	-3.1203	-0.02563	-6.73	0.000	3.0

S = 0.1255 R-sq = 25.3% R-adj = 20.9%

Source	DF	SS	MS	F	P
Regression	6	8.1626	1.3604	93.63	0.000
Residual Error	20	0.0524	0.0026		
Total	26	8.2150			

[그림2] 예측 모델에 대한 통계량

4.1.2 최적화 모델

2차 항의 예측모델을 만들었으므로 Quadratic Programming에 의해서 최적수준을 결정할 수 있다. 이때, 공정인자의 제어 범위, 또는 제어 희망 범위를 제약조건으로 둔다. 식(6)은 중요 공정인자의 조업조건을 찾기 위한 최적화 모델이다.

$$\begin{aligned} \text{Minimize} & \quad y = -10.6 - 0.233x_{13} - 3.12x_{16} + 0.0133x_{95} + 0.842x_7x_{52} \\ & \quad + 0.0861x_7x_{142} + 0.170x_{15}x_{52} \\ \text{subject to} & \quad 0.046 \leq x_7 \leq 0.0564 \\ & \quad 3.07 \leq x_{13} \leq 3.24 \\ & \quad 0.25 \leq x_{15} \leq 0.33 \\ & \quad 0.025 \leq x_{16} \leq 0.031 \\ & \quad 14 \leq x_{52} \leq 18 \\ & \quad 841 \leq x_{95} \leq 882 \\ & \quad -55 \leq x_{142} \leq -20 \end{aligned} \quad (6)$$

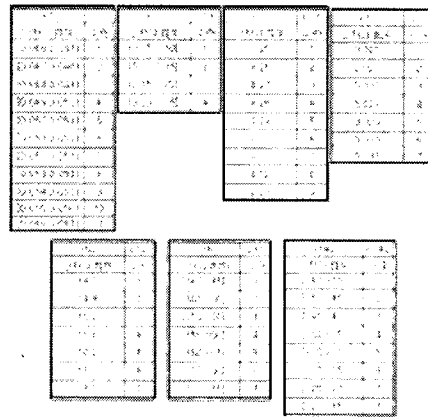
식(6)에 의해서 (x₇, x₁₃, x₁₅, x₁₆, x₅₂, x₉₅, x₁₄₂) = (0.046, 3.24, 0.25, 0.031, 14, 841, -55)을 조업조건으로 얻었다. 여기서는 모두 식(6)에서 제약조건 상한 또는 하한값과 일치한다.

4.2 비함수적 방법에 의한 조업조건

4.2.1 중요 공정인자의 조업수준 결정

현장에서 관리되는 공정인자는 제어형태에 따라서 이산형 제어와 연속형 제어로 분류된다. 본 사례에서 x₁₅, x₁₆, x₅₂가 이산형 제어인자에 해당하고 x₇, x₁₃, x₉₅, x₁₄₂는 연속형 제어인자에 해당한다.

공정인자의 조업수준은 전문가의 지식을 반영하여 이산형 제어인자의 경우, 제어단위를 각 조업수준으로 정하고, 연속형 제어인자의 경우는 인자의 제어 범위를 조업수준으로 정한다. [그림3]은 중요 공정인자의 조업수준을 결정한 결과를 보여준다.



[그림3] 중요 인자의 조업수준

4.2.2 조업수준의 최적 조합 결정

중요인자의 수준의 조합은 1,086,624 개 (= 11×4×9×7×7×7×8)이다. 각 조합에 해당 하는 품질인자 y의 평균값을 비교함으로써 최적 조합을 결정하고 유의성을 평가하는 것은 불가능하다. 본 논문에서는 단순회귀의 R²를 이용하여 조업 수준을 결정할 공정인자를 먼저 선택하고, 최적 조업조건을 ANOVA를 이용하여 결정한다. 일단, 한 개의 중요 공정인자에 대해서 조업조건이 결정되면 결정된 조업조건을 만족하는 조업 데이터에 대해서 위의 과정을 반복함으로써 모든 중요 공정인자의 조업조건을 결정하게 된다.

중요 공정인자는 X₇, X₁₃, X₁₅, X₁₆, X₅₂, X₉₅, X₁₄₂ 인데, 다음은 본 사례에서 공정인자의 조업수준을 결정한 실제 과정을 처음 두 개의 인자에 대해서 보여준다.

● X₁₄₂의 선택

7개의 중요 공정인자 중에서 단순회귀의 R²가 제일 큰 X₁₄₂를 가장 먼저 조업조건을 결정할 중요 공정인자로 선택한다. [표2]는 단순회

귀 분석 후, R²를 정리한 것이다.

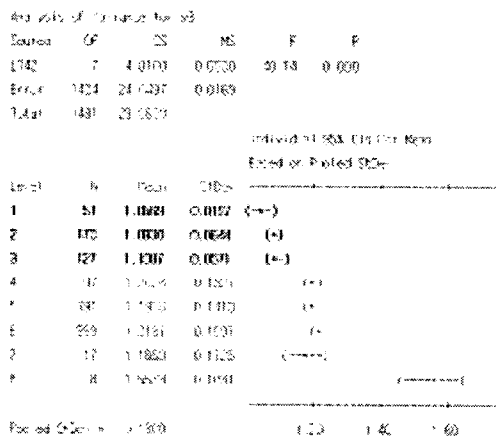
[표2] 조업데이터의 단순회귀 R²

	X7	X13	X15	X16	X52	X95	X142
R ² (%)	1.24	0.43	7.38	6.02	4.33	11.6	13.5

● X142의 조업조건 결정

X142의 One-Way ANOVA 분석 결과에 의해서 X142의 조업조건을 1, 2, 3으로 결정한다. [그림4]는 X142의 One-Way ANOVA 분석 결과이다. X142가 수준1일 때에 품질인자의 평균이 가장 낮으며, 수준 1은 수준2와 3과 유의한 차이가 없음을 확인할 수 있다.

One-way ANOVA of versus L142



[그림4] X142의 One-Way ANOVA 결과

● L142 ∈ {1, 2, 3}일 때, X95의 선택

조업 데이터에서 X142가 수준 1, 2, 3을 만족할 때, X95가 단순회귀의 R²가 가장 크기 때문에 두 번째로 조업 조건을 결정할 인자로 선택한다. [표3]은 단순회귀 분석 후, R²를 정리한 것이다.

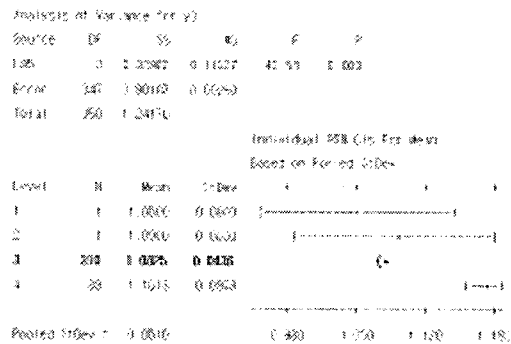
[표3] L142 ∈ {1, 2, 3}을 만족하는 조업 데이터에서의 단순회귀 R²

	X7	X13	X15	X16	X52	X95
R ² (%)	2.97	1.90	1.58	0.57	0.02	20.45

● X95의 조업조건 결정

X95의 ANOVA 분석 결과, X95의 조업조건을 3으로 결정한다. [그림5]에서 수준3과 수준4는 확연히 수준3이 좋은 품질을 나타냄을 알 수 있고, 수준1과 수준2는 해당 관측치가 1개 밖에 없으므로 조업조건에서 제외한다.

One-way ANOVA of versus X95



[그림5] X142의 조업조건이 결정된 후에 X95의 One-Way ANOVA 결과

4.2.3 결과 요약

중요 공정인자의 최적 조합이 결정되었으면, 공정인자의 수준을 실제 조업조건으로 환원해야 한다. 이때, [그림3]의 정의를 이용할 수도 있으나 보다 엄격한 범위를 제시하기 위해 실제 데이터를 살펴보는 것을 추천한다. [표4]는 비합수적 방법에 의한 조합조건의 결과를 보여준다.

[표4] 비합수적 방법에 의한 조업조건 결과

공정인자	수준	조업조건
X7	6,7,8,9	[0.051,0.054]
X13	1,2,3	[3.08,3.20]
X15	4,5,6	0.28, 0.29, 0.30
X16	2,3,4,5	0.026, 0.027, 0.028, 0.029
X52	2,3,4	14.9, 15.0, 15.1
X95	3	854, 855
X142	1,2,3	[-55,-42]

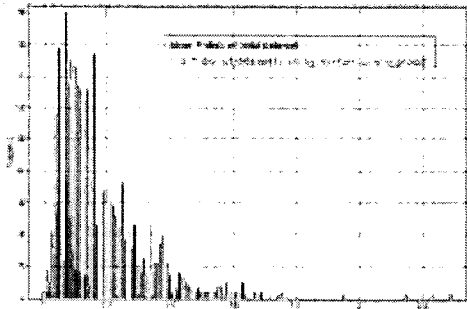
[표5]는 1432개의 관측치를 갖는 조업 데이터에서 비합수적 방법에 의한 조업조건을 만족하는 품질인자의 수준에 대한 통계량을 요약한다. Y는 조업데이터의 품질인자이고 Y*는 조업데이터에서 조업조건을 만족하는 품질인자를 나타낸다. [표5]의 통계치로부터 비합수적 방법에 의해서 찾아진 조업조건을 준수하면 제품의 품질이 개선됨을 기대할 수 있다.

[표5] 품질인자의 비교

	개수	평균	표준편차	최소	최대
Y	1432	1.1815	0.1420	1	2.29
Y*	203	1.0740	0.0426	1.01	1.34

[그림6]은 Y와 Y*의 분포를 나타내는데, Y*의 분포가 Y보다 확연히 개선되었음을 눈으로

확인할 수 있다.



[그림6] 품질인자의 분포

4.3 함수적 방법과 비함수적 방법의 비교

[표6]은 조업조건 결정방법에 대한 장단점을 요약한다. 예측모델의 정확도가 높고 실험이 어렵지 않은 경우에는 함수적 방법에 의한 조업조건 결정이 타당하고 예측모델의 정확도가 낮고 실험이 어려운 경우에는 비함수적 방법이 타당한 것으로 판단된다.

[표6] 조업조건 결정방법의 비교

	조업조건 결정방법	
	함수적 방법	비함수적 방법
장점	과거보다 더 좋은 결과를 낳는 조업 조건을 찾을 수 있다.	정확한 예측 모델이 필요 없고 조업데이터 안에서 해를 찾기 때문에 해에 대한 신뢰도가 높다.
단점	정확도가 높은 예측 모델을 만들기가 힘들고 해를 검증하기 위한 추가 실험의 부담이 크다.	해를 조업 데이터에서 찾기 때문에 공정을 개선할 수 있는 한계가 있다.

5. 결론

본 논문은 현장의 센서에 의해서 자동으로 수집되는 조업 데이터를 바탕으로 품질인자에 영향을 미치는 중요 공정인자를 선택하고 각 중요 공정인자의 조업 조건을 결정하는 절차를 개발하였다. 세부적으로는 예측모델을 사용하지 않고 직접 조업조건을 찾는 방법을 개발하였으며 실제 사례에 적용한 구체적인 예를 제시하였다. 본 논문은 품질개선을 위한 조업조건을 찾는 산업현장의 좋은 예가 될 것으로 기대되며 제안된 절차의 각 모듈에 다른

기법을 적용해 봄으로써 보다 좋은 결과를 도출할 수 있을 것으로 사료된다.

감사의 글

이 논문은 2003년도 IMT 2000 사업 (과제번호: 00015993)에 의하여 지원 되었음

참고문헌

- [1] Geladi, Paul and Bruce R. Kowalski (1986), "Partial Least-Squares Regression: A Tutorial", *Analytica Chimica Acta*, 185, 1-17
- [2] Hastie, Trevor et al. (2001), *The Elements of Statistical Learning: Data Mining Inference and Prediction*, New York: Springer
- [3] Montgomery, D.C. et al. (2001), *Introduction to Linear Regression Analysis*, 3rd, Wiley