

# GMDH 방법을 이용한 FMS의 성능 예측 방안의 연구

## A GMDH-type Performance Modeling for FMS with Unreliable RAM and LCC

황 홍 석  
Heung-Suk Hwang

### ABSTRACT

통합생산시스템에서의 고장, 정비 및 가용 도는 매우 중요한 역할을 한다. 시스템 설계시의 RAM 파라미터의 결정은 시스템의 성능과 소요비용 및 구성(System Configuration) 등에 크게 영향을 미친다. 이러한 시스템관련 요소의 숫자가 많거나 불확실할 경우는 시스템의 성능예측이 매우 복잡하게된다. 이러한 시스템의 성능(Performance) 평가를 위하여 발견적 방법인 GMDH(Group Method Data Handling) Type Modeling 방법을 이용하여 FMS의 성능 평가를 시도하였다. RAM 및 기계작업시간의 Data로부터 시스템성능의 척도로서 단위 사이클 기간동안의 생산률, 시스템내의 총 흐름시간, 각 작업장의 기계의 RAM 및 LCC 등을 고려하였다. GMDH 알고리즘의 계산을 위한 프로그램을 개발하고, 이를 L형 Bracket 제조시스템의 성능 예측에 시험 적용하였다. 본 Modeling에 의한 시스템의 성능예측 방법은 입출력 자료의 처리과정을 개선할 경우 FMS계획 및 운영 단계에서 성능평가에 매우 유용하게 활용될 수 있을 것으로 본다.

### 1. 서 론

통합생산시스템에서의 고장, 정비 및 가용도(RAM)는 매우 중요한 역할을 한다. 통합생산시스템 설계시의 RAM 파라미터의 결정은 시스템의 성능, 비용요소 및 시스템구성(System Configuration) 등에 크게 영향을 준다. 이러한 요소들이 불확실할 경우는 문제가 더욱 복잡해진다. 특히 FMS의 초기설계단계에서의 RAM관련자료의 분석은 대부분의 경우에 누락되어 있다(Grady 와 Menon(1986)).

FMS시스템에서의 유연성(Flexibility)을 위해서는 먼저 여건 변화에 쉽게 대처할 수 있어야 하고, 이러한 변화를 예측하고 이의 변화에 적응하거나 피해갈 수 있어야 하며, 이러한 예측 및 대처하는 방법을 자동화시스템으로 이루어져야 한다(Ranky(1983)). 통합자동생산시스템의 설계문제는 이에 관련된 많은 파라메타들이 필요하며, 이중 중요한 것을 들면 다음과 같다:

- 작업품목의 도착률(Arrival Rate)
- 제품의 생산 조합(Production Mix)
- 작업지점(Work Center) 및 가공시간(Processing Time)
- 작업지점간의 거리(Location)
- 물류시스템(MHS)
- 작업지점의 여유대기규모(Buffer Size)
- 시스템의 총 제조능력(Total Part Capacity)
- 시스템 및 서부시스템의 신뢰도(MTBF)
- 시스템의 정비능력(MTBR)
- 시스템의 작업주기

본 연구에서는 이러한 시스템성능에 관련된 많은 파라메타로 부터 통합생산시스템의 성능을 예측하는 방법으로 우선 시스템 및 서부시스템의 신뢰도와 정비도의 파라메타와 각 설비의 비용요소로부터 순기비용(LCC) 및 가용도(Availability)를 산정하고 이들을 이용하여 시스템 성과 산정을 위한 모델링을 시도하였다. 계산을 위하여 GMDH-Type Modeling 방법을 사용하고 GMDH알고리즘의 과정을 전산화하고 Sample 시스템의 자료를 적용하여 그 결과를 검증하였다.

일반적으로 통합자동시스템의 성능평가기 관련 시스템의 신뢰도 및 정비도와 가공공정시간의 자료가 불확실할 경우의 모델링에 관한 연구결과들을 보면, 1)분석적인 대기행렬모델(Analytical Queueing Models) 과 2)시뮬레이션모델(Simulation Model)로 분류할 수 있다. 본 연구에서는 이들을 활용하고, 통계자료분석방법을 사용하였다.

통합생산 시스템에서의 Simulation문제는 복잡한 시스템의 특성과 많은 Parameter들을 포함할 수 있는 특성을 가지고 있기 때문에 널리 활용되고 있다. 즉 많은 작업장(Workstation)의 설비, 가공제품의 배합, 가공시간, 고장 및 정비관련 요소들을 포함한 Modeling이 가능하기 때문이다. 이를 위한 많은 Simulation Package들이 개발되어 사용되고 있으며 이중에서 FMS 시스템성능 평가에 쉽게 응용될 수 있는 것으로 SIMAN, SLAM, FACTOR/AIM(Priksker(1992))등을 들 수 있다. 이러한 시뮬레이션 방법과 함께 통계적인 방법중의 한가지인 GMDH방법(Group Method of Data Handling)(Farlow(1984))이다. 이 방법은 최초 Ivakhnenko(1968)에 의해서 개발되었으며, 최근에 Chao 및 Ferreira(1986)는 제조공정문제에 GMDH 방법을 응용하였으며 그림 3에서와 같이 다 계층 문제로 접근하였다.

이 방법은 각 계층에서의 시스템 성과를 적절한 기준으로 평가하여 다음 단계를 지날 때마다 매우 복잡한 Model을 가지게 된다. 즉 마지막 단계의 Model은 각 단계마다 2개의 변수로서 이루어진 수식이 되며 각 변수는 최하위 단계의 2개 변수의 함수를 이룬다. 처음단계에서 한 개의 출력변수  $Y_0$ 와  $m$ 개의 입력변수  $x_1, x_2, \dots, x_m$ 가 주어지면, 각 입력 변수의 쌍,  $x_i$  와  $x_j$ 에 대하여 회귀분석식을 다음과 같이 설정한다.

$$Y_0 = A + Bx_i + Cx_j + Dx_i^2 + Ex_j^2 + Fx_i x_j$$

가능한 모든 독립변수의 조합  $(m(m-1)/2)$ 가지에 대하여 각각 회귀분석하고 일정기준 이상의 결과를 다음 단계의 입력변수로 사용하여 최적 해를 얻을 때까지 되풀이한다. 기존의 단일단계의 회귀분석과 비교하면 GMDH방법은 다음과 같은 몇 가지 이점을 갖는다.

1) GMDH 방법은 적은 DATA로서도 많은 수의 Parameter를 포함하는 모델을 만들 수 있다. 예를 들면  $m$ 개의 입력변수를 포함하는  $P$ 차 Regression Polynomial은 아마도  $(1/P!)(\prod_{i=1}^P(m+i))$ 항을 포함하는 다차 방정식이 될 것이다.  $m=10$  및  $P=8$ 인 경우는 약 43,758개항을 가지는 복잡한 수식이 된다. 일반적인 회귀분석 방법으로 이 각 항의 계수를 구하기란 매우 어렵다. 반면 GMDH알고리즘(Chao 와 Ferreira(1988))에 의하면 6개의 계수만 구하는 문제로 해결될 수 있다.

2) GMDH알고리즘은 매 단계마다 일정한 기준에 미달하는 변수는 다음 단계에서 제외되므로 불필요한 중복성을 줄일 수 있다.

3) GMDH에서와 같이 다단계 방법을 통한 다차 다항식(High Degree Polynomial)의 경우, 일반적인 회귀 분석 방법으로는 적절한 Model식을 구하기가 어렵다.

GMDH 방법에서는 Output에서 매 단계마다 영향이 큰 변수를 포함하고 영향이 적은 변수를 제외시키는 과정이 있어 이러한 문제는 쉽게 극복될 수 있다. 이러한 GMDH 기법을 확장한 많은 연구들이 있으며 (Farlow(1984), Ikeda(1976), Ivakhnenko(1968)(1971)(1988)), 본 연구에는 이를 불확실한 DATA하에서의 FMS의 성능평가에 확장하여 응용하였다.

## 2. GMDH-Type 모델링

### 2.1 통합생산시스템의 모델링

본 연구에서 고려한 통합생산시스템은 다음과 같은 특징을 갖은 시스템이다.

- DATA들을 사전에 관측이 어렵다.
- 이들 DATA들은 많은 수의 변수에 관련된다.
- 관측가능한 변수들을 분석하는 적절한 수리적 분석방법을 용이하게 활용하기 어렵다.
- 가능한 DATA는 실제 작업장의 실제가공 문제에 적용하기가 용이하지 않다.

먼저 주어진 문제를 정의하면 아래와 같다.

$Y$ 를 통합생산시스템의 성능척도(Performance Measure),  $X$ 를 성능에 관련된 파라미터를 표시하는 집합이라고 하자. 여기서  $X$ 의 요소로서 신뢰도(MTBF), 정비도(MTR) 및 순기비용(LCC) 등을 뜻하며  $E[X]$ ,  $E[Y]$ 를 각각  $X$ ,  $Y$ 의 기대값으로 두면, 다음과 같은 관계식을 생각할 수 있다.

$$E(Y_1) = f(E(X_1), E(X_2), \dots, E(X_m), X_{m+1}, \dots, X_n)$$

$$E(Y_2) = f(E(X_1), E(X_2), \dots, E(X_m), X_{m+1}, \dots, X_n)$$

⋮

$$E(Y_k) = f(E(X_1), E(X_2), \dots, E(X_m), X_{m+1}, \dots, X_n)$$

위의 수식들로부터  $X_i$ 의 각 파라미터를 구하기 위하여 일반적 회귀 분석 방법을 사용할 수도 있으나 앞의 항에서와 같이 실제문제에서는 각 독립변수  $X_i$ 의 규모에 따라 극히 제한된 범위에서 가능할 것이다. 본 연구에서는 많은 변수들을 포함하고 다차 다항식으로 표시되는 문제를 GMDH방법으로 Modeling 하기 위한 과정을 그림 1 에서와 같이 요약하였다.

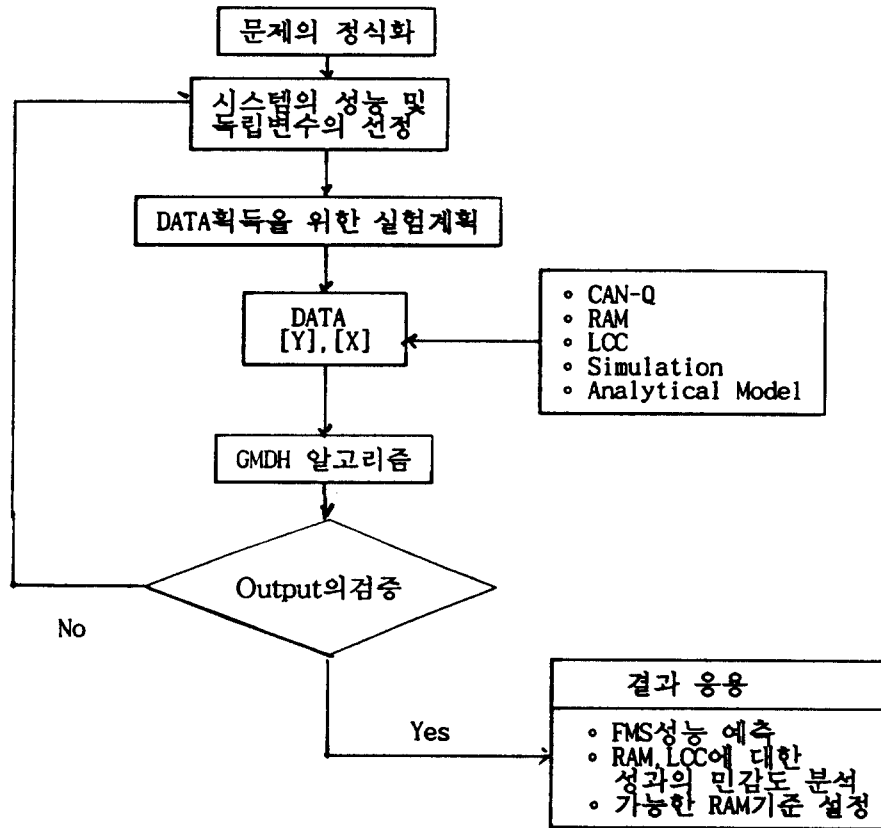


그림 1. GMDH-type Modeling의 흐름도

#### ① 문제의 정식화

통합생산시스템의 분석문제를 정식화하기 위하여 GMDH TYPE의 형태로 문제를 정의해야 한다. 우선 통합생산시스템의 설계단계에서의 문제에서 획득가능한 신뢰도, 정비도, 및 가용도 등이 사전에 가용할 자료가 없을 경우 우선 GMDH-Type 모델링을 위한 문제의 정식화가 되어야 한다. 다음으로 운영단계에서는 소요되는 각종 운영 파라미터의 변경, 예방정비 및 생산제품의 비용분석이 필요시 된다.

#### ② 시스템의 성능 및 독립변수의 선정

본 연구에서 고려된 통합생산시스템의 성능에 관련된 독립변수로 신뢰도(MTBF), 정비도(MTTR), 비용(LCC)등을 고려하였으며 이들의 DATA를 위하여 시뮬레이션 및 분석적인 방법들을 사용하였다.

#### ③ DATA 획득을 위한 실험계획(Experimental Design)

이 단계에서는 GMDH 모델링에 필요한 DATA의 획득을 위하여 변수의 조합 및 DATA 획득을 위한 실험계획(Experimental Design)을 하는 단계이다. 앞 단계에서 선정된 각 독립변수의 파라메타 값들로부터 DATA 획득을 위한 계획을 하는 단계이다. 획득한 DATA SET의 수와 각 DATA의 조합 등이 이루어진다.

#### ④ DATA 획득

이 단계에서는 위에서 계획된 DATA를 위하여 통합생산시스템의 실제 DATA를 획득하거나, 어려운 경우에는 시뮬레이션 방법에 의하여 획득한다. 이를 위하여 본 연구에서는 먼저 관련 독립변수 X의 각 파라메타를 조합하는 프로그램(DATGEN)과 CAN-Q, 및 RAM 프로그램과 SIMAN, FACTOR/AIM등의 통합 생산 시스템의 전용 시뮬레이터를 사용하여 구하였다.

⑤ 위에서와 같이 획득한 DATA를 GMDH 알고리즘에 적용시킨다. 이는 4개의 단계에 따라 이루어지며 다음항에서 구체적인 설명이 이루어질 것이다.

- ⑥ GMDH 알고리즘에 의하여 구해진 MODEL을 새로운 DATA SET(Checking Data)을 적용하여 예측된 결과와 실제결과와의 오차를 검증한다. 이 결과 만족한 결과를 얻지 못한 경우 ②의 단계부터 되풀이한다.
- ⑦ 구해진 예측모델을 산정 하여 통합생산시스템의 성능예측, 민감도분석, 및 가능한 RAM 기준설정 등에 응용한다.

## 2.2 DATA 획득

본 Modeling에 필요한 Data는 가능할 경우 실제 시스템으로부터 관측된 DATA를 사용해야 하나, 실제 Data의 획득이 어려울 경우 시뮬레이션 및 기타 방법에 의하여 DATA를 획득한다. 이 경우 Data 획득을 위한 실험계획(Experimental Design)방법으로 본 연구에서는LATIN HYPERCUBE DESIGN (Mckay와 Backman(1979)) 방법을 사용하였다. 우선 DATA 실험계획표를 표 1에서와 같이 독립변수 집합 X와 성과를 나타내는 Y 변수를 필요한 DATA 수 만큼 획득하도록 계획한다.

표 1. Data의 실험 계획표(예:)

구분	$X_1$	$X_2$	$X_3 \dots X_n$	$Y_1$	$Y_2 \dots Y_k$
1	32.0	2.0	29.0 . . 1.9	7.1	21.0 . . 5.0
2	31.0	2.5	28.0 . . 1.8	7.2	23.0 . . 4.9
3	33.0	2.1	28.5 . . 2.0	6.9	22.0 . . 5.1
		:			:
		:			:

본 실험계획법은 선정된 각 변수(X 및 Y)의 상한과 하한을 먼저 선정하고 이 구간 내에서 균등분포를 가정하고 Data를 산출하는 과정이다. Latin Hypercube Design 방법에 의하여 산출된 X의 파라미터들로부터 그림 2에서와 같이 CAN-Q, RAM 모델 및 시뮬레이션 방법으로부터 시스템의 생산률, 가동률 및 기타 X 및 Y 파라미터로 추가되어야 할 변수들을 보완하여, GMDH 알고리즘을 응용하기 위한 Data를 산출한다. 위에서와 같은 Data 산출과정을 그림으로 표시하면 그림 2와 같다.

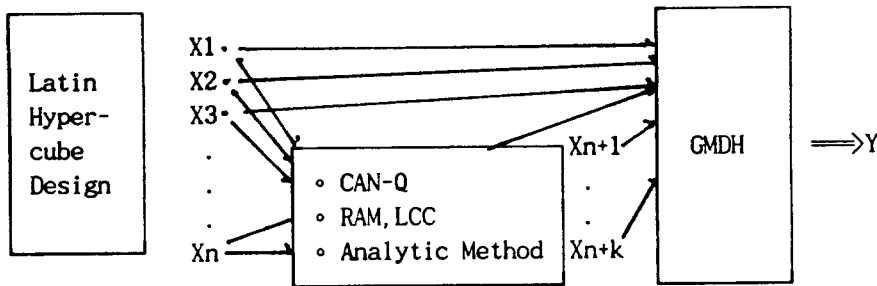


그림 2. Data 산출과정

## 3. GMDH 알고리즘

GMDH 알고리즘은 Ivakhnenko(1968)에 의해서 개발된 이후 시스템 컨트롤 분야에 널리 응용되어 왔으며 이후 환경 및 경제성 분석 분야의 Modeling에 널리 응용되고 있다. GMDH 알고리즘은 그림 3에서와 같이 다단계(Multi-layered)적 구조를 이루고 있으며 각 단계에서 입력 Data로부터 비선형 변형을 하며 이는 통상 2개의 선정된 변수의 2차 다항식을 이룬다.

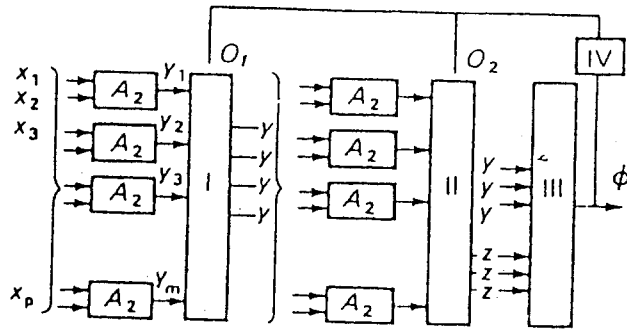


그림 3. GMDH 다단계구조(Chao 및 Ferreira(1986)에서 인용)

$Y_0$ 가 시스템의 성능변수이고,  $m$ 개의 입력 변수를  $X_i$ ,  $i=1, m$ 일 경우, 임의의 2개의 변수  $X_i$ , 및  $X_j$ 의 조합으로 부터 다음과 같은 회귀분석식을 사용하였다.

$$Y_0 = A + BX_i + CX_j + DX_i^2 + EX_j^2 + EX_i X_j$$

여기서 입력 변수의 조합의 경우수는  $m(m-1)/2$ 가 되며  $m(m-1)/2$ 번의 회귀분석으로부터의  $Y_0$ 를 얻는다. 위의 결과로부터  $K$ 개의 다항식을 선정(선정기준을 회귀분석 우열기준에 의거함)하여 다음 단계의 입력 Data로 사용한다. 두번째 단계에서 얻은 2차 다항식,  $Y_0$ 는 최초독립변수의 4차식이 된다.

최종단계의 예측은 매우 복잡한 다차원의 다항식이 되며 이를 "Ivakhnenko Polynomial" (Chao(1988))로 알려져 있다. 예를 들면 10개의 초기 독립변수로부터 10단계 후의 얻은 다항식은최초 독립변수의 256차 다항식이 되며 이 식을 이용한 회귀분석은 매우 복잡하게 된다. 그러나 GMDH 알고리즘은 매 단계마다 2차 다항식이 되므로 우선 이러한 복잡한 문제는 피할 수 있다. 통합 자동생산시스템의 성능을 표시하는 변수  $Y$ 의 예측 Model의 일반식을 다음과 같이 표시 할 수 있다.

$$Y = f(x_1, x_2, x_3, \dots, x_m)$$

위에서 얻은 가용한 Data로부터  $f(x_1, x_2, x_3, \dots, x_m)$ 의 각 예측 파라미터를 구하는 문제를 위한 GMDH 알고리즘을 다음과 같이 4개의 단계로 요약하였다:

#### 단계1 : Data 구분

가용한 Data를 2개의 부분으로 구분하여, 파라미터 예측을 위한 Data부분(Training Data Set)과 오차를 검증하는 Data부분(Checking Data Set)으로 구분한다.

#### 단계2 : 새로운 변수의 설정

초기변수로부터 각각 2개의 변수를 선정하여 2차 다항식을 이루는 예측 Model을 선정하고 Training Data Set으로부터 예측 식을 구한다.

예 :  $f(x_1, x_2, x_3)$ 인 경우  $x_1, x_2$ ;  $x_1, x_3$ ;  $x_2, x_3$ 의 3가지 조합이 가능하다.

이 단계에서는 최초의  $m$ 개의 변수로 구성되는 다차 다항식 대신 2개의 변수로 구성되는 2차다항식의 예측 Model을 사용한다.

#### 단계 3: 기여도가 적은 변수의 선별

이는 Checking Data Set로부터 2단계에서 구해진 예측 결과를 검증하여 다음단계의 입력변수로 활용하는 단계이다. 이를 선별하기 위하여 다음과 같은 선별식을 사용하였다.

$$r_j^2 = \frac{\sum_{i=nt+1}^n (y_i - Z_{ij})^2}{\sum_{i=nt+1}^n y_i^2} \quad j = 1, 2, \dots, \binom{m}{2}$$

여기서  $i=nt+1$ ,  $n$  까지의 합은 Data의 Checking Set의 결과를 뜻하며  $Z_{ij}$ 는  $i$ 번째 Data들로 부터  $j$

번째 변수 조합으로부터 회귀분석한 결과치이다. 이 단계에서 얻은 예측 값은  $r_j < R$  인 것만 다음단계의 입력변수로 활용된다(  $R$  는  $r_j$  의 허용한계치이며 사전에 사용자가 정의한다).

**단계 4 : 최적화 검증**

단계 3으로부터  $r_j$  의 최소값(RMIN)을 구하여 이를 전 단계의 RMIN값과 비교하여 개선되지 않을 경우 전 단계의 예측식을 사용하고 회귀분석을 끝내며, 그렇지 않을 경우 다음단계(Layer)로 계속된다. RMIN과 예측반복횟수 간의 관계를 표시하면 그림 4와 같다.

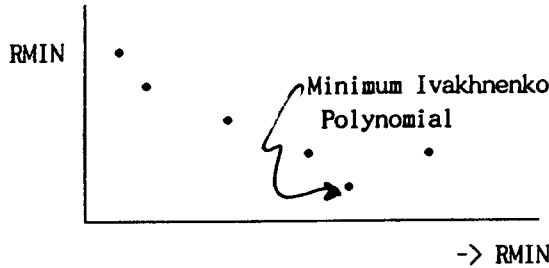


그림 4. RMIN과 예측회수

이 단계에서 최종 예측식은 초기단계의 입력변수  $x_1, x_2, x_3, \dots, x_m$  의 다차 다항식으로 최적 "Ivakhnenko Polynomial"이라고 하며 다음과 같이 표시된다.

$$Y_0 = a + \sum_{i=1}^m b_i x_i + \sum_{i=1}^m \sum_{j=1}^m c_{ij} x_i x_j + \dots + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k$$

**4. GMDH-Type Modeling의 응용**

**4.1 시스템개요**

위의 GMDH-Type Modeling을 6개의 작업센터(Work station)로 구성되는 통합생산시스템의 성능 예측 문제에 응용하였다. 예시된 시스템은 그림 5 에서와 같이 2대의 가공기계(W/S #1)가 있고, 세척 작업장(W/S #2)에서 세척된 가공품을 건조시키고 도색작업장(W/C #3)에서 도색한 후 검사작업장(W/S #4)에서

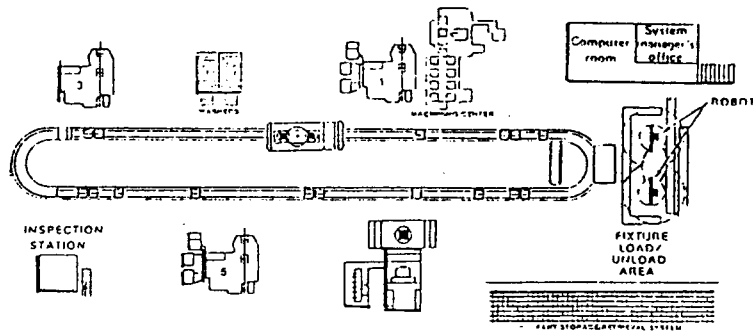


그림 5. Sample 시스템의 Layout

검사결과에 따라 약 20%정도는 수정작업장(W/S #5)으로 가서 수정후 마지막 포장작업(W/S #6)에서 포장후 적하장(L/UL)을 거쳐 시스템을 떠나게 되며 나머지 약 80%정도는 바로 포장작업(W/S #6)으로 가서 포장후 적하장(L/UL)을 거쳐서 시스템을 떠난다. 예시된 시스템의 운영을 위하여 1대의 AGV를 사용하고 있으며 3가지 Type의 가공품을 가공한다. 본 시스템의 가공 흐름도를 그림 6에서와 같이 표시하였다.

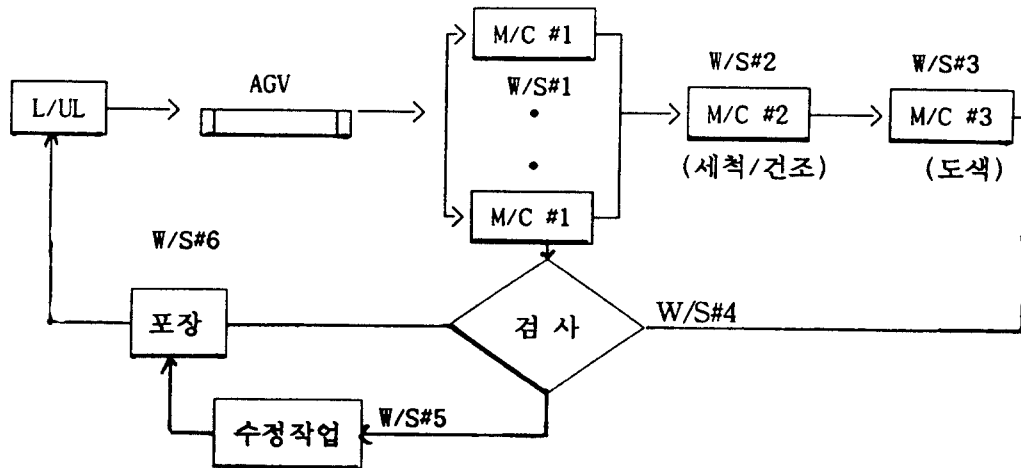


그림 6. 시스템 흐름도(예:)

#### 4.2 시스템 모델링

본 시스템의 성능예측을 위하여 고려할 주요변수를 요약하면 다음과 같다.

◦ 시스템 성능측정(Y:)

$Y_1$  : 시스템 생산률(Production Rate)

$Y_2$  : 시스템의 순기비용(LCC)

◦ 시스템 입력 변수(X:)

$X_1$  : W/S #1의 고장시간(MTTF 1)

$X_2$  : W/S #2의 고장시간(MTTF 2)

$X_3$  : W/S #3의 고장시간(MTTF 3)

$X_4$  : W/S #3의 정비시간(MTTR 3)

$X_5$  : W/S #5의 고장시간(MTTF 5)

$X_6$  : W/S #6의 정비시간(MTTR 6)

$X_7$  : SYS의 가용시간당 LCC

먼저 위에서 정의한 시스템의 입력 변수( $x_1, x_2, x_3, \dots, x_7$ ) 및 출력변수( $y_1, y_2$ )의 Data는 실측자료 활용이 어려운 경우를 가정하고 다음과 같이 획득하였다.

$x_1, x_2, x_3, \dots, x_6$ 까지의 MTTF 및 MTTR 관련 자료는 다음과 같이 상이한 균등분포로 가정하고 Latin Hyper Cube Design방법으로 30개 Data 포인트를 구하였으며,  $x_7$  및  $y_1$ 은 CAN-Q 및 RAM프로그램으로부터 각각 30개 Data포인트를 산출하였다. 예시된 시스템의 출력변수로 고려한  $Y_i$ 는 CAN-Q 및 RAM프로그램으로부터 산출하였으며 다음 2가지를 고려하였다.

$Y_1$  : 생산률(Production Rate)

$Y_2$  : 단위생산제품당 순기비용(LCC)

각 W/S에서의 MTTF 및 MTTR의 상하한값을 다음과 같이 가정하였다.

W/S #1	MTTF	25.0	,	44.0
	MTTR	2.5	,	3.5
W/S #2	MTTF	35.0	,	44.0
	MTTR	2.5	,	3.5
W/S #3	MTTF	26.0	,	39.0
	MTTR	2.1	,	3.9
W/S #4	MTTF	30.0	,	39.0
	MTTR	2.5	,	3.5
W/S #5	MTTF	40.0	,	49.0
	MTTR	2.5	,	3.5
W/S #6	MTTF	45.0	,	54.0
	MTTR	3.1	,	3.9

각 W/S에서의 MTF 및 MTR (MTFi, MTRi)의 Data를 위하여 Latin Hyper Cube의 Model의 FORTRAN프로그램에 의하여 각각 30개 Data 포인트를 산출하고, CAN-Q 및 RAM프로그램을 이용하여 시스템 생산 능력, 단위 생산품당 LCC 및 LCC의 Data를 산출하고 시스템 출력변수의 Data로 사용하였으며 이를 표 2에서와 같이 표시하였다.

표 2. Sample 시스템의 GMDH 입력 Data

BRACKET PROBLEM GMDH INPUT DATA:

7	30	15	0	1	0.0				
4.091349	25.630000	2.700000	32.930000	2.850000	46.800000	3.820000	31.907900		
3.864084	29.430000	2.770000	39.000000	2.900000	48.900000	3.450000	34.244130		
3.972037	26.900000	3.270000	28.600000	2.310000	49.800000	3.390000	33.156600		
3.806935	38.930000	2.570000	30.770000	2.710000	50.700000	3.690000	34.875280		
4.004428	34.500000	3.100000	29.030000	2.660000	51.600000	3.420000	32.676120		
3.991446	30.070000	3.400000	30.330000	2.610000	49.500000	3.290000	33.143740		
4.122840	42.730000	2.830000	33.800000	2.230000	51.900000	3.550000	31.470220		
3.945699	36.400000	2.870000	35.530000	2.290000	50.100000	3.150000	33.353930		
3.985143	41.470000	2.900000	27.300000	2.450000	53.400000	3.710000	33.106410		
3.971757	35.130000	2.730000	28.170000	2.420000	53.100000	3.900000	33.271950		
3.995379	28.170000	3.030000	37.270000	2.820000	45.900000	3.230000	32.858940		
3.852815	31.330000	3.170000	29.900000	2.210000	47.400000	3.630000	34.142850		
4.129625	35.770000	3.500000	26.430000	2.150000	54.000000	3.310000	31.596660		
3.965621	37.030000	2.600000	34.230000	2.390000	51.300000	3.580000	33.280540		
4.034701	30.700000	3.470000	38.130000	2.500000	50.400000	3.130000	32.731540		
4.003622	42.100000	3.300000	35.970000	2.870000	52.500000	3.210000	32.675940		
3.955661	32.600000	3.130000	36.400000	2.790000	52.800000	3.260000	33.342780		
3.836138	38.300000	2.970000	36.830000	2.470000	48.600000	3.530000	34.602850		
4.012189	28.800000	3.430000	31.200000	2.630000	45.600000	3.850000	32.846250		
3.917648	33.870000	2.930000	31.630000	2.580000	47.700000	3.500000	33.500740		
3.994479	37.670000	3.330000	33.370000	2.130000	53.700000	3.660000	32.912080		
3.904152	26.270000	2.670000	37.700000	2.690000	49.200000	3.870000	33.782830		
4.488213	40.830000	2.630000	29.470000	2.340000	48.000000	3.740000	29.513720		
4.050287	27.530000	3.070000	38.570000	2.180000	48.300000	3.770000	32.310750		
3.928362	31.970000	2.800000	27.730000	2.770000	46.200000	3.610000	33.624810		
4.128178	43.370000	3.200000	35.100000	2.530000	47.100000	3.790000	31.813640		
3.944860	44.000000	3.000000	32.500000	2.260000	46.500000	3.470000	33.456490		
3.944860	33.230000	2.530000	26.870000	2.740000	52.200000	3.370000	33.456490		
4.128178	39.570000	3.230000	34.670000	2.370000	51.000000	3.180000	31.813640		
3.944860	40.200000	3.370000	32.070000	2.550000	45.300000	3.340000	33.456490		

위의 GMDH 알고리즘에 의한 각 Level에서의 계산을 위하여 PC Level에서 사용할 수 있도록 FORTRAN으로 개발하였다. 위의 Data를 GMDH 알고리즘에 따라 2개 부분, Training data Set와 Checking Data Set로 나누어 Training Data Set으로부터 각 Level에서의 예측 파라메타를 구하고 Checking Data Set으로부터 예측 Data와 실 Data간의 오차를 산정 하였다. 회귀분석 결과의 적합성 판단을 위하여 다음과 같은  $r_j$  을 산정 한다.

$$r_j = \left( \frac{\sum_{i=nt+1}^n (y_i - Z_{ij})^2}{\sum_{i=nt+1}^n y_i^2} \right)^{\frac{1}{2}}, \quad j = 1, 2, \dots, \binom{m}{2}$$

여기서  $i=nt+1$ ,  $n$  까지의 Data는 Checking Data Set을 뜻하며  $Z_{ij}$  는  $i$ 번째 Data로부터  $j$ 번째 변수 조합 경우의 회귀분석 결과이다.

위의 각 Level에서  $r_j$  값 중에서 가장 적은 값을 해당 Level의 RMIN 값으로 한다. 다음 Level에 사용



할 Data는 해당 Level에서  $r_j$  가 일정한 값 R (사용자가 정의함)보다 작은 경우의 출력 자료만으로 구성된 Data Set이 후속 Level의 입력 Data로 사용된다.

RMIN값이 후속 단계의 RMIN값 보다 적게 될 경우에 그 Level의 예측값을 최적으로 하고 GMDH과정을 끝내도록 하였다.

### 4.3 모델의 응용 결과

표 3. 예제 문제의 GMDH 출력 결과

```

-----
LEVEL NUMBER= 1
NO. VARIABLES SAVED = 7
RMIN VALUE (SUMMED OVER CHECKING SET)= .66629E-04
MULTIPLE CORRELATION (SUMMED OVER TRAINING SET) = .997206
-----

GMDH CONVERGED AFTER 1 GENERATION(S)
MULTIPLE CORRELATION (SUMMED OVER TRAINING SET) = .997206
-----

```

CASE NO.	OBSERVED VALUE	ESTIMATE	ERROR	PERCENT ERROR
1	.40913490E+01	.41067380E+01	-.15388490E-01	-.37612260E+00
2	.38640840E+01	.38552540E+01	.88298320E-02	.22851030E+00
3	.39720370E+01	.39674930E+01	.45437810E-02	.11439420E+00
4	.38069350E+01	.38170640E+01	-.10128740E-01	-.26606010E+00
5	.40044280E+01	.40195920E+01	-.15164380E-01	-.37869020E+00
6	.39914460E+01	.39676650E+01	.23780820E-01	.59579470E+00
7	.41228400E+01	.41297610E+01	-.69212910E-02	-.16787680E+00
8	.39456990E+01	.39498380E+01	-.41391850E-02	-.10490370E+00
9	.39851430E+01	.39948090E+01	-.96664430E-02	-.24256200E+00
10	.39717570E+01	.39789690E+01	-.72124000E-02	-.18159220E+00
11	.39953790E+01	.39986110E+01	-.32322410E-02	-.80899480E-01
12	.38528150E+01	.38551620E+01	-.23474690E-02	-.60928680E-01
13	.41296250E+01	.41185040E+01	.11120800E-01	.26929310E+00
14	.39656210E+01	.39638660E+01	.17547610E-02	.44249330E-01
15	.40347010E+01	.40105650E+01	.24135590E-01	.59820020E+00
16	.40036220E+01	.40237460E+01	-.20124440E-01	-.50265570E+00
17	.39556610E+01	.39705000E+01	-.14838700E-01	-.37512560E+00
18	.38361380E+01	.38170900E+01	.19047980E-01	.49654040E+00
19	.40121890E+01	.40011990E+01	.10989670E-01	.27390700E+00
20	.39176480E+01	.39252680E+01	-.76198580E-02	-.19450080E+00
21	.39944790E+01	.40125900E+01	-.18111470E-01	-.45341250E+00
22	.39041520E+01	.39029580E+01	.11937620E-02	.30576730E-01
23	.44882130E+01	.43762880E+01	.11192510E+00	.24937570E+01
24	.40502870E+01	.40538190E+01	-.35324100E-02	-.87213810E-01
25	.39283620E+01	.39085520E+01	.19809480E-01	.50426830E+00
26	.41281780E+01	.41156690E+01	.12508870E-01	.30301190E+00
27	.39448600E+01	.39281200E+01	.16739610E-01	.42433970E+00
28	.39448600E+01	.39556720E+01	-.10811570E-01	-.27406720E+00
29	.41281780E+01	.40989690E+01	.29209140E-01	.70755520E+00
30	.39448600E+01	.39282860E+01	.16574140E-01	.42014530E+00

IVAKIINENKO POLYNOMIAL

```

Y = A + B*U + C*V + D*U*U + E*V*V + F*U*V
A = .1944E+02
B = -.2431E+00
C = -.4762E+00
D = .7205E-03
E = .1739E-02
F = .5286E-02

```

U = X( 5)      V = X( 7 )

GMDH 프로그램의 출력 결과를 표 3에서와 같이 표시하였다. 이는 사용자가 예측 Model의 식을 입력할 필요가 없으며 각 Level에서의 예측 시마다 프로그램 내에서 자동적으로 주어지도록 하였으며 시스템의 성능과 관련한 변수가 많을 경우는, GMDH 알고리즘은 2차 다항식으로 계산되므로 매우 간편하게 구할 수 있다. 표 3의 출력 결과에서 보면, 예제의 예측 Model은 다음과 같이 주어진다.

$$Y_0 = 0.0019 - 0.2431X_5 - 0.4762X_7 + 0.00072X_5^2 + 0.1739X_7^2 + 0.0052X_5X_7$$

본 Model의 예측오차가 매우 적음을 알 수 있으며, 이는 실제 획득 가능한 Data의 가용이 비슷한 결과를 얻을 수 있을 것으로 본다.

## 5. 결론

위에서와 같이 GMDH 알고리즘을 이용한 통합생산시스템의 성능예측결과는 실측치와의 오차가 크지 않으며, 복잡한 시스템의 경우 자동 모델링 기능이 있으므로 2차 다항식으로 간편하게 접근할 수 있다. 보다 더 발전되어야 할 분야로서 GMDH 알고리즘의 필요한 Data 획득 과정부터 전 과정을 프로그램에 포함되도록 보완되고 입출력변수의 정의 및 구분 등이 매우 중요한 분야로 생각된다.

여기서 제시한 GMDH-Type Modeling 방법은 기존의 방법들을 이용하여 필요한 입력 Data로 신뢰도, 정비도, 가용도 및 비용

등을 통합생산시스템과 연결하여 방법론과 실용적인 측면에서 예제를 통하여 그 결과를 검증하였다. 자동통합생산시스템의 성능 예측, FMS의 설계 및 운영단계에서의 활동을 고려한 Modeling이 가능하다.

즉:

- 1) 설계 단계
  - 수용 가능한 장비의 신뢰도 및 정비도(MTBF, MTTR)를 고려하여
    - RAM 자료와 성능과의 비교
    - RAM 조건간의 Trade off가 가능하다.
  - 시스템성능 관련 관계식을 예측함으로써 설계 조정에 유용하게 활용할 수 있다.
- 2) 운영단계
  - 통합생산시스템에 대한 장비의 열화로 인한 RAM조건의 변화에 따른 시스템의 성능을 예측할 수 있으며,
  - 가공품(Part)의 조합 및 공정조정에 따른 시스템의 성능을 예측가능하다.
  - 다양한 정비정책(PM)을 연구 가능하다.

본 연구에서 개발된 GMDH 프로그램은 입출력 절차를 개선할 경우 통합생산시스템의 불확실한 Data환경에서 매우 유용하게 활용 수 있으며 그 예측 오차도 크게 줄일 수 있을 것으로 본다.

## 참고문헌

- [1] Chao, P. Y., Ferreira, P. M., and Liu, C. R., "Applications of GMDH-type Modeling in Manufacturing", *Journal of Manufacturing Systems*, 7, 241-52, 1988.
- [2] Farlow, S. J., *Self-Organizing Methods in Modeling GMDH-type Algorithms*, Marcel Dekker Inc., New York, 1984.
- [3] Grady, P. O., "A Concise Review of Flexible Manufacturing Systems and FMS Literature", *Computers in Industry* Vol. 7, pp 115-167, 1986.
- [4] Ikeda, S., Ochiai, M., and Sawaragi, Y., "Sequential GMDH Algorithm and its Application to River Flow Prediction", *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6, 473-79, 1976.
- [5] Ivakhnenko, A. G., "The Group Method of Data Handling, a Rival to Stochastic Approximation", *Soviet Automatic Control*, 13, 43-55, 1968.6] Ivakhnenko, A. G., "Group Handling of Data in Identification of the Static Characteristic of a Multi-Extremal Plant", *Soviet Automatic Control*, 14, 30-37, 1969.
- [7] Ivakhnenko, A. G., "Polynomial Theory of Complex Systems", *IEEE Transactions on Systems, Man, Machine Cybernetics*, SMC-1, 364-78, 1971.
- [8] Ivakhnenko, A. G., "Sorting Methods for Modeling and Clusterization (Survey of GMDH Papers for the Years 1983-1988) The Present", *Soviet Journal of Automation and Information Sciences*, 4, 1-13, 1988.
- [9] McKay, M. D. and Beckman, R. J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, 21, 239-45, 1979.
- [10] Prisker Corporation, "FACTOR/AIM : Finite Capacity Management", 1993.
- [11] Ranky. P., *The Design and Operation of FMS*, North-Holland, New York, 1983.