

신경망을 이용한 동적 수율 개선 모형

정현철* · 강창욱** · 강해운*[†]

*한양대학교 산업공학과

**한양대학교 정보경영공학과

Dynamic Yield Improvement Model Using Neural Networks

Hyun-Chul Jung* · Chang-Wook Kang** · Hae-Woon Kang*[†]

*Dept. of Industrial Engineering, Hanyang University

**Dept. of Information and Industrial Engineering, Hanyang University

Yield is a very important measure that can express simply for productivity and performance of company. So, yield is used widely in many industries nowadays. With the development of the information technology and online based real-time process monitoring technology, many industries operate the production lines that are developed into automation system. In these production lines, the product structures are very complexity and variety. So, there are many multi-variate processes that need to be monitored with many quality characteristics and associated process variables at the same time. These situations have made it possible to obtain super-large manufacturing process data sets. However, there are many difficulties with finding the cause of process variation or useful information in the high capacity database. In order to solve this problem, neural networks technique is a favorite technique that predicts the yield of process for process control. This paper uses a neural networks technique for improvement and maintenance of yield in manufacturing process. The purpose of this paper is to model the prediction of a sub process that has much effect to improve yields in total manufacturing process and the prediction of adjustment values of this sub process. These informations feedback into the process and the process is adjusted. Also, we show that the proposed model is useful to the manufacturing process through the case study.

Keywords : Neural Networks, Yield Improvement, Process Control

1. 서론

기업에서는 경쟁력을 표현하고자 생산주기(cycle time), 재작업율(rework rate), 공정변동(process variation), 공정재고(work-in-process), 수율(yield) 등 공정 성과와 품질 성능 지표들을 관리하고 있다. 다양한 성능지표 중 투입 수에 대한 적합품 비율을 의미하는 수율은 생산성, 수익성, 업체의 성과 측면에서 쉽게 표현이 가능한 척도이다. 특히, 수율은 기업 경쟁력의 근간이라 할 수 있으며, 반

도체 산업, LCD, PDP 등 전자산업, 자동차산업, 부품산업 등에서 중요한 척도로 사용되고 있다.

기업의 생산라인들은 정보기술의 발전과 더불어 온라인 기반의 실시간으로 공정을 관리하는 자동화 시스템으로 발전하여 운영되고 있다. 이러한 자동화 시스템은 인터넷과 같은 온라인 환경의 빠른 보급으로 빠르게 발전하고 운영되고 있다. 이러한 환경에서는 공정에서 측정되는 대용량의 공정 데이터를 실시간으로 수집하고 저장한다. 이렇게 수집된 공정 데이터는 단지 이상 발생

여부를 판단하는 자료로 사용되어 왔으며 공정개선이나 품질개선에 관련된 유용한 정보를 찾아내기 위한 자료로는 활용되어지지 않고 있다[1].

제품 구조가 복잡해지고 다양해지면서, 생산라인들은 하나의 특성치보다는 여러 개의 연관된 특성치들과 공정 변수들을 동시에 관리해야 하는 다변량 공정들이 많아지게 되었다. 반도체와 같은 공정 산업에서는 한 공정에서 수많은 특성치가 발생하고 있으며, 제품의 품질은 이런 많은 특성치가 결합하여 결정된다. 이와 같은 다변량 공정은 다양한 공정변수가 있고 이들이 상호작용하면서 수율과 비선형의 복잡한 관계가 있기 때문에 수율 관리뿐만 아니라 수율 저하의 원인을 신속하고 정확하게 파악해서 개선하기가 더욱 더 어렵다[2].

이러한 다변량 공정을 관리하기 위하여 X^2 통계량 혹은 Hotelling T^2 통계량을 이용한 다변량 통계적 공정 관리도가 개발되었다[6]. 그러나 이러한 기법들은 다변량 정규분포 기반으로 각 특성치들이 다변량 정규분포를 가정하고 있다. 또한 특성치들은 서로 간의 상관관계를 알고 있거나 과거 데이터로부터 예측 가능하다고 가정하고 있다. 그리고 다변량 관리도를 적용 할 경우에는 공정변수와 데이터가 많아지면 계산량이 많아지게 되고, 그 해석도 매우 어렵다[7].

이와 같은 문제를 해결하기 위하여, 신경망과 같은 지능적인 도구를 이용한 공정 수율 예측과 방법들에 대한 연구들이 활발하게 진행되고 있다. Timothy[8]는 고주파 담금질 공정에서 신경망을 이용하여 정도에 영향을 주는 공정변수를 예측하여 공정에 적용하였다. 이장희[3]는 반도체와 LCD 제조 공정과 같이 많은 요인들이 비선형적으로 상호 작용하면서 수율에 영향을 주는 공정 시스템에서 수율을 예측하고 관리 하기위해 데이터 마이닝 도구들의 혼합 적용하여 정확히 예측할 수 있는 방법을 제안하였다.

실제 자동화 생산라인에서는 정확한 수율 예측 능력 향상뿐만 아니라 실시간으로 운용이 가능한 수율 개선, 관리 방법 및 모형 개발 또한 매우 중요하다. 본 연구에서는 실시간 및 온라인 기반의 자동화 시스템 환경에서 수율 개선에 영향을 많이 주는 공정과 조정 값을 예측하고 이를 실시간으로 공정에 피드백(feedback)하여 동적으로 수율을 개선 및 유지하는 모형을 제안하고자 한다.

2. 신경망을 이용한 동적 수율 개선 방법

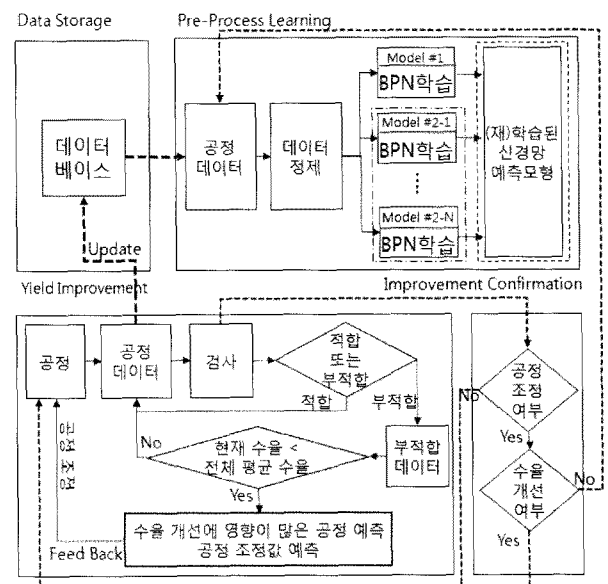
2.1 개요

일반적으로 기업체에서는 제조 공정의 수율을 향상시

키거나 현재 수율을 유지하려고 다양한 공정의 특성을 알아내어 모형화하며, 원하는 상태로 이들 공정을 관리하는 방법에 관심을 두고 있다. 효과적인 수율관리를 하려면 품질 특성치에 영향을 주는 요인들과의 관계를 체계적으로 정리해야한다.

실제 현장에서는 생산기술의 발전으로 공정의 변수가 복잡해지고, 생산라인의 자동화, 정보 기술의 발전으로 공정 데이터들이 방대하게 수집된다. 이렇게 얻어진 공정 데이터는 변화범위가 상대적으로 작고, 시계열적 종속성이 있는 등 여러 가지 요인 때문에 통계적 기법이나 엔지니어의 경험적 분석만으로는 파악하지 못하는 수율 저하 요인이 다수 존재한다[4]. 특히, 수많은 공정 변수와 대응량의 데이터는 모든 변수를 분석하기에는 너무 많은 시간이 소요되어 실제적으로 일부 변수들만 분석이 되고 있다. 그리고 이러한 공정들은 엔지니어가 미처 인식하지 못하고 지나치는 공정 이상들이 존재하고 있으며, 단위 공정별 품질개선 활동은 활발히 이루어지고 있으나 전체공정을 고려한 종합적 수율 개선 노력은 부족한 상황이다[2].

이러한 환경 하에서는 수율에 영향을 주는 요인들과의 관계를 찾는 것도 중요하지만, 온라인에서 실시간으로 데이터가 축적이 되고 자동화된 생산라인 환경에서는 직접적으로 수율에 영향을 많이 주는 공정을 찾아 개선하는 것이 더 효과적이고 중요하다. 따라서 본 연구에서는 신경망을 이용하여 수율 개선에 영향을 많이 주는 공정을 예측하고, 동시에 그 공정을 개선하기 위한 변수 조정값을 예측한다. 그리고 예측된 정보를 공정에 피드백하고 조정하여 공정의 수율을 동적으로 개



<그림 1> 제안 모형

선하는 모형을 제시한다. 앞의 <그림 1>은 본 연구에서 제안 하는 동적수율 개선모형의 전체 흐름도이다.

2.2 데이터 저장과 사전 학습 단계

데이터 저장과 사전 학습 단계는 공정의 데이터를 데이터베이스에 저장하고, 수율 개선 모형에서 이용하는 데이터를 정제하여 예측력이 우수한 신경망 예측 모형을 학습하는 단계이다. 이 단계에서는 공정에서 측정된 데이터를 데이터베이스에 저장하고, 신경망 예측 모형을 학습하기에 앞서 실제 데이터에 대하여 충분한 사전 처리를 거친다. 그리고 수율 개선에 필요한 신경망 예측 모형을 학습한다. 학습한 신경망 예측 모형은 예측 능력을 향상시키려고 주기적으로 재학습한다. 데이터 저장과 사전 학습 단계는 신경망의 예측 능력이 제안하는 모형의 성능과 밀접한 관계가 있기 때문에 매우 중요한 단계이다.

학습은 신경망 학습 알고리즘 중 가장 일반적으로 사용되는 역전파(backpropagation) 알고리즘을 이용하여 수율 개선에 필요한 신경망 예측 모형을 학습한다. 이 학습 단계에서는 학습된 신경망 예측 모형의 예측능력을 향상시키고자 실시간으로 저장되는 데이터를 이용하여 주기적으로 재학습을 한다. 공정 데이터의 입력 변수들은 비선형조합(non-linear combination)을 통해 결과 값을 제공하므로 예측력이 우수한 신경망 기법을 사용한다. 신경망은 연결강도에 의하여 상호 연결되어진 노드들의 집합이다. 노드들이 서로 독립적으로 작동되므로 병렬 처리와 분산 처리가 가능한 구조를 갖고 있으며, 이에 따라 빠른 정보처리와 외부잡음에 강건한 특성을 가지

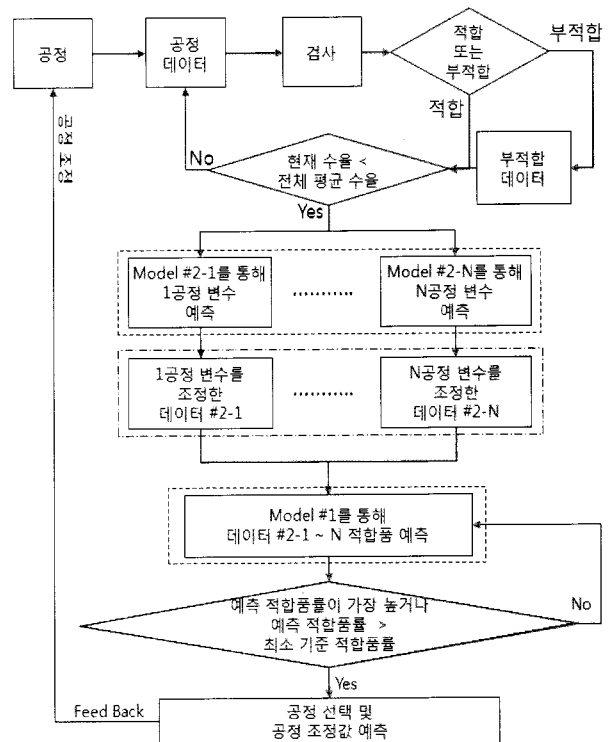
<표 1> 학습을 위한 신경망 예측 모형

수율 개선에 영향을 많이 주는 공정을 예측하기 위한 모형 (Model #1)	공정 별 변수를 예측하기 위한 모형 (Model #2-1~N)
<ul style="list-style-type: none"> 공정의 여러 변수 값을 통해 품질 특성(적합품 또는 부적합품)을 예측 	<ul style="list-style-type: none"> k 공정이 포함하고 있는 변수 예측 ($1 \leq k \leq N, N$: 전체 부분 공정의 총 개수)
<ul style="list-style-type: none"> 입력변수(Input Layer) : - 모든 공정이 포함하고 있는 변수 	<ul style="list-style-type: none"> 입력변수(Input Layer) : - k 공정을 제외한 나머지 공정이 포함하고 있는 변수 - 품질 특성치(적합품 또는 부적합품)
<ul style="list-style-type: none"> 은닉층(Hidden Layer) 수 : - 사용자 정의 	<ul style="list-style-type: none"> 은닉층(Hidden Layer) 수 : - 사용자 정의
<ul style="list-style-type: none"> 예측결과(출력변수 : Output Layer) : - 품질 특성치(적합품 또는 부적합품) 	<ul style="list-style-type: none"> 예측결과(출력변수 : Output Layer) : - k 공정이 포함하고 있는 변수
<ul style="list-style-type: none"> 학습률(η) : 사용자 정의 	<ul style="list-style-type: none"> 학습률(η) : 사용자 정의
<ul style="list-style-type: none"> 모멘트 값(α) : 사용자 정의 	<ul style="list-style-type: none"> 모멘트값(α) : 사용자 정의

고 있다[5]. 신경망의 학습 알고리즘 중 역전파 알고리즘을 이용하여 학습된, 신경망 예측 모형 구성은 아래의 <표 1>과 같다.

2.3 수율 개선 단계

수율 개선 단계에서는 수율을 모니터링하여 수율 저하시, 학습된 신경망 예측 모형을 이용하여 수율 개선에 영향이 많은 공정과 변수 조정값을 예측하고, 이 정보를 공정에 피드백하여 조정하는 단계이다. 이 단계에서는 사전 학습 단계에서 학습한 신경망 예측 모형을 이용하여 품질 특성치가 적합품이 되게 하는 부분 공정별 변수값을 예측한다. 그리고 부분 공정을 변수 예측값으로 조정된 것에 대한 수율 개선 효과를 예측하여 수율 개선에 영향이 많은 부분 공정을 예측한다. 또한 수율 개선에 영향이 많은 공정의 변수 조정값을 신경망 예측 모형으로 예측한다. 이렇게 예측한 정보를 실시간으로 공정에 피드백하고 조정하여 공정의 수율을 동적으로 개선한다. 아래 <그림 2>는 본 연구의 제안 모형 중 수율을 개선하는 단계의 전체 흐름도이다. 수율을 관리하는 부분, 수율 개선에 영향이 많은 공정과 변수 조정 값을 예측하는 부분, 예측한 정보를 공정에 피드백한 후 조정하는 부분들로 구성되어 있다. 수율 개선 단계는 아래와 같은 세부절차들을 통해 수행된다.



<그림 2> 수율 개선 단계의 세부 흐름도

- 세부절차 1 : 측정된 공정변수와 품질 특성치 구성된 공정 데이터에서 품질 특성치가 부적합품인 데이터를 수집한다.
- 세부절차 2 : 수율을 관리한다. 현재 공정 수율이 공정의 전체 평균 수율 보다 낮은 경우, 수율 개선에 영향을 많이 주는 공정과 변수 조정 값 예측을 실시한다.
- 세부절차 3 : 부적합품 데이터를 이용하여, 품질 특성이 적합품이 되도록 하는 공정 별 변수 값들을 Model #2-1~N로 예측한다. 그리고 공정 별 변수 예측 값으로 조정된 데이터 집합(데이터 #2-1~N)을 구성한다. 조정된 데이터 집합(데이터 #2-1~N)을 Model #1로 적합품을 예측한다.
- 세부절차 4 : 세부절차 3을 통해 예측한 공정 별 적합품률을 비교하여 수율 개선에 영향이 많은 공정을 선택한다. 예측 적합품률을 비교하여 부분 공정을 선택 할 때는 공정 별 예측 적합품률 중 가장 높게 예측된 부분 공정을 선택하거나 최소 기준 적합품률을 설정하여 최소 기준 적합품률보다 높은 예측 적합품률을 보이는 부분 공정을 선택한다. 여기서, 최소 기준 적합품률은 예측 적합품률이 낮은 공정을 조정할 경우에는 수율 개선에 의미가 없는 것으로 판단되어, 이를 사전에 판단하고 수율 개선에 높은 효과를 도출하기 위한 최소의 기준치를 말한다. 최소 기준 적합품률은 엔지니어의 판단과 공정의 상태에 따라 결정 한다. 본 연구에서는 예측 적합품률을 비교하여 공정 별 예측 적합품률이 가장 높게 예측된 부분공정을 수율 개선에 영향을 많이 주는 부분 공정으로 예측한다.
- 세부절차 5 : 세부절차 4에서 수율 개선에 영향이 많은 것으로 예측한 공정(m)의 변수 조정 값을 Model #2-m 으로 예측한다(N : 전체 부분 공정 총 개수, $1 \leq m \leq N$).
- 세부절차 6 : 세부절차 4에서 수율 개선에 영향이 많은 것으로 예측한 공정(m)에 세부절차 5에서 예측한 변수 조정값을 피드백하여 조정한다.

2.4 수율 개선 후 확인 단계

수율 개선 후 확인 단계는 실제 공정이 조정 되었을 때, 수율이 공정 조정이 되기 이전보다 향상 되었는지 확인하는 단계이다. 이 단계는 수율 개선 단계에서 수율 개선에 영향이 많은 공정과 변수 조정 값은 신경망을 이용하여 예측한 정보이므로 실제 수율이 개선되었는지 확인하는 단계가 필요하다.

수율 개선 후 확인 단계는 수율 개선 여부를 확인하

여 조정된 공정의 유지 여부를 판단한다. 제안 모형을 이용하여 공정을 조정한 후 공정의 수율이 공정의 전체 평균 수율로 향상되어 유지되는 경우에는 예측한 변수 조정 값으로 공정을 유지하여 운영하면 된다. 그러나 공정 수율이 향상되지 못하여 공정의 전체 평균 수율보다 낮은 경우에는 제안 모형에서 사용되는 신경망 예측 모형인 Model #1, Model #2-1~N을 데이터베이스에 추가로 저장된 공정 데이터를 이용하여 즉시 재학습하여 예측 모형의 정확도를 높인다. 재학습을 통해 Model #1, Model #2-1~N의 예측 능력을 향상시킨 후, 수율 개선에 영향이 많은 공정과 변수 조정값을 다시 반복 예측한다.

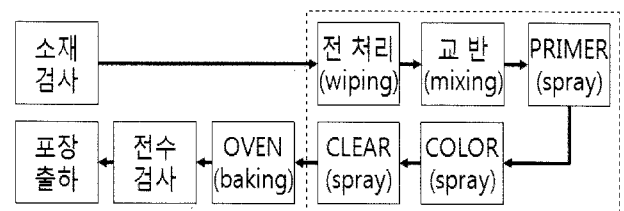
3. 적용 예제

본 연구에서 제안하는 신경망을 이용한 동적 수율 개선 모형의 이해를 돕기 위해 실제 현장에서 측정된 데이터를 적용하고자 한다. 예제에 사용한 데이터는 자동차용 내·외장품 생산 업체인 A기업의 자동차 범퍼 도장 공정라인에서 부적합품이 다량 발생한 이상 상태(out of control) 시점의 학습용 데이터 840개, 평가용 데이터 390개 등 총 1,230개 샘플 데이터를 1회 수집하였으며, 신경망 학습을 위한 도구로는 SPSS 사의 클레멘타인(Clementine) 10.0을 이용하였다.

3.1 공정 소개

본 예제에서 적용된 공정은 자동차의 외관을 보호하는 범퍼를 도장하는 공정이다. 자동차 범퍼는 다양한 기후와 환경조건 속에 오랜 기간 사용되기 때문에 자동차 범퍼 도장 공정은 높은 수준의 품질과 기술이 필요하다. 특히, 범퍼는 우레탄 계열의 제품이어서 열 반응에도 쉽게 치수가 줄어들거나 늘어날 수 있다. 이와 같은 특징 때문에 자동차 범퍼 도장 공정은 일반 도장 공정과는 비교할 수 없을 정도로 복잡한 공정을 갖고 있다. 자동차 범퍼 도장 공정은 <그림 3>과 같이 소재 검사부터 포장 출하까지의 과정으로 구성되어 있다. 그

<그림 3> 공정 전체 흐름도와 주요 관리 공정



중 자동차 범퍼 도장 공정은 전처리 공정에서 클리어 공정까지로 총 다섯 공정을 주요공정으로 관리한다.

자동차 범퍼 도장 공정에는 작업조건의 이상, 작업자의 작업 미숙, 설비 이상 등 여러 가지 원인들에 의해 부적합이 발생한다. 여기서, 작업자의 작업 미숙, 설비 이상과 같은 원인들에 의한 발생하는 부적합품은 매우 유동적이고 데이터 수집이 용이하지 않다. 본 연구에서는 데이터 수집이 용이한 주요 관리 공정의 작업조건에 의한 부적합품을 고려한다. 따라서 공정 변수로는 주요공정의 온도, 습도, 교반 공정에서 원액에 약품을 섞어주는 교반 시간을 포함한 9개의 변수를 선택하였다. 실험에 사용된 변수와 변수가 의미한 내용은 <표 2>와 같다. 여기서, P_i 는 주요관리공정 중 i 번째 공정을 의미하고, C_{ij} 는 P_i 공정의 j 번째 변수를 의미한다. 는 첫째 공정인 전처리 공정을 의미하고 는 공정인 전처리 공정 내의 온도를 의미하고 한다.

<표 2> 공정변수 정의

공정	각 공정 내 변수들	변수
전처리 공정(P_1)	온도(°C) 습도(%)	C_{11} C_{12}
교반 공정(P_2)	교반 시간(초)	C_{21}
프라이머 공정(P_3)	온도(°C) 습도(%)	C_{31} C_{32}
컬러 공정(P_4)	온도(°C) 습도(%)	C_{41} C_{42}
클리어 공정(P_5)	온도(°C) 습도(%)	C_{51} C_{52}
품질 특성치	적합품(0) 또는 부적합품(1)	Y

3.2 데이터 저장과 사전 학습 단계

본 적용 예제에 사용한 데이터들은 A기업의 자동차 범퍼 도장 공정라인 데이터 1,230개로 학습용 데이터 840개, 평가용 데이터 390개로 구성되어 있다. 자동차 범퍼 도장 공정에서 수집한 데이터를 이용하여 분석에 용이한 데이터로 재구성하기 위해 신경망 예측 모형 학습에 앞서 데이터를 사전 전처리를 하였다.

먼저, 수집된 데이터에 결측치와 이상치가 있는지 확인하였다. 그 결과 수집된 데이터 1,230개에서는 결측치와 이상치가 없는 것을 확인하였다. 따라서 수집된 데이터 1,230개를 신경망 예측 모형 학습과 평가에 사용하였다. 신경망 예측 모형은 범위가 0에서 1사이 값인 데이터를 학습했을 때 최적의 성능을 제공한다. 따라서 신경망 예측 모형을 학습하기 전 최적의 성능의 신경망

예측 모형을 학습하고자 학습용 데이터 840개의 모든 독립변수와 종속 변수를 0에서 1사이 범위의 값으로 정규화하였다. 아래의 <표 3>은 학습용 데이터 840개의 평균, 표준편차, 최대 값, 최소 값을 나타내고 있다. 종속변수인 공정 변수는 (실제 값-최소 값)/(최대 값-최소 값)으로 변환하였다. 즉, P_1 공정의 C_{11} 값인 26은 $(26-20)/(29-20) = 0.667$ 값으로 변환하였다. 종속변수인 품질 특성치는 이항형(binary-type) 변수이므로 부적합품을 1로 적합품을 0으로 변환하였다.

<표 3> 공정변수 내 통계량

공정	변수	개수	평균	표준편차	최대값	최소값
P_1 공정	C_{11}	840	25.780	1.816	29	20
	C_{12}	840	63.419	4.788	77	50
P_2 공정	C_{21}	840	30.390	1.635	34	24
P_3 공정	C_{31}	840	26.291	1.344	29	21
	C_{32}	840	64.517	5.856	79	50
P_4 공정	C_{41}	840	26.985	1.729	32	21
	C_{42}	840	64.742	6.196	87	42
P_5 공정	C_{51}	840	25.798	1.632	29	20
	C_{52}	840	62.531	5.622	76	45

<표 4> 적용 예제에서의 신경망 예측 모형

Model #1	Model (#2-1~#2-5)
<ul style="list-style-type: none"> ◦ 입력변수(Input Layer) : - 총 9개 : $C_{11} \sim C_{52}$ ◦ 은닉층(Hidden Layer) 수 : - 총 4개 	<ul style="list-style-type: none"> ◦ 입력변수(Input Layer) : - 각 공정($P_1 \sim P_5$)을 제외한 나머지 공정이 포함하고 있는 변수들 + 품질 특성치(Y)(즉, 8-9개) ◦ 은닉층(Hidden Layer) 수 : - 총 4개
<ul style="list-style-type: none"> ◦ 출력변수(Output Layer) : - 총 1개 : Y 	<ul style="list-style-type: none"> ◦ 출력변수(Output Layer) : - 각 공정($P_1 \sim P_5$) 공정이 포함하고 있는 변수 개수 (즉, 1-2개)
◦ 학습률(η) : 0.3	◦ 학습률(η) : 0.3
◦ 모멘트값(α) : 0.9	◦ 모멘트값(α) : 0.9

이와 같이 정규화한 학습용 데이터를 역전파 알고리즘을 통해 수율 개선에 필요한 신경망 예측 모형을 학습하였다. 이와 관련한 신경망 예측 모형의 모델 구성은 위 <표 4>와 같다. 본 연구의 학습 모델에서 사용되는 학습률(η)과 모멘트값(α)은 클레멘타인(Clementine)에서 제공되는 기본값인 0.3, 0.9를 각각 적용하였으며, 은닉층(hidden layer)의 수는 입력변수(input layer)의 절반정도인 4개를 각각의 모델들에 사용하였다.

학습용 데이터 840개를 이용하여 신경망 예측 모형을 학습한 후, 학습한 신경망 예측 모형을 평가하였다. 학습한 신경망 예측 모형에 평가용 자료 390개에 대해서 품질 특성치인 적합품과 부적합품을 예측한 결과 387개를 정확하게 예측하여 99.231%의 예측 정확도를 나타냈다. 즉, 390개 중에서 실제 적합품 223개 중 220개를 정확하게 적합품으로 예측하였고, 부적합품 167개를 모두 정확하게 부적합품으로 예측한 것을 의미한다. 3개는 적합품을 부적합품으로 잘못 예측한(오분류) 결과값이 존재하였다. 이 결과들을 통하여 신경망 모형의 예측이 높은 정확도를 나타내고 있음을 파악하였다.

3.3 수율 개선 단계

본 적용 예제에서 사용되는 데이터의 공정은 수율 관리 시 30개의 데이터가 누적 될 때마다 수율을 지속적으로 측정한다. 그리고 본 공정의 전체 평균 수율은 60%이다. 생산된 제품 1,230개의 데이터에서 현재 공정의 수율을 측정하였다. 생산된 제품 1,200개 데이터의 수율은 60.106%로 부적합품이 479개였으나 30개의 제품이 더 생산되면서, 30개 중에서 18개가 부적합품으로 되어 현재 공정의 수율이 1,230개 중 497개의 부적합품이 생산되었으므로 59.593%으로 내려갔다. 따라서 현재 공정 수율이 공정의 전체 평균 수율인 60%보다 낮기 때문에 수율을 개선하고자 수율 개선에 영향을 많이 주는 공정과 변수 조정값을 예측을 실시하였다.

우선, 생산된 제품 1,230개 데이터 중 부적합품 데이터 497개를 수집하였다. 이렇게 수집한 부적합품 데이터에서 품질 특성이 적합품이 되도록 하는 공정 별 변수값을 학습한 신경망 예측 모형 Model #2-1~5로 예측하였다. 적합품이 되도록 하는 P_1 공정의 변수값은 P_2 공정부터 P_5 공정까지 실제데이터 변수인 $C_{21}, C_{31}, C_{32}, C_{41}, C_{42}, C_{51}, C_{52}$ 의 값과 적합품을 의미하는 Y 변수의 0값을 입력변수로 하여 Model #2-1을 통해 예측하였다. $C_{21}, C_{31}, C_{32}, C_{41}, C_{42}, C_{51}, C_{52}$ 가 31, 25, 66, 24, 56, 22, 52인 경우에 적합품이 되게 하는 C_{11}, C_{12} 은 27.673, 53.614로 예측되었다. 그리고 나머지 P_2 공정부터 P_5 공정의 변수도 부적합품 데이터와 공정 별 변수 예측모형인 Model #2-2~5를 이용하여 적합품이 되도록 하는 각각의 변수를 동일한 과정으로 예측하였다. 공정 별 변수 예측값을 이용하여 공정 별 조정된 데이터 집합(데이터 #2-1~5)을 구성하였다. <표 5>은 부적합품 데이터를 Model #2-1로 예측한 P_1 공정 변수값을 이용하여 구성된 조정된 데이터 집합(데이터 #2-1)이다.

<표 5> P_1 공정 변수 예측값을 이용한 조정된 데이터 (데이터 #2-1 예)

부적합품 번호	변수 예측값		실제 변수값							
	P_1 공정		P_2 공정	P_3 공정		P_4 공정		P_5 공정		
	C_{11}	C_{12}	C_{21}	C_{31}	C_{32}	C_{41}	C_{42}	C_{51}	C_{52}	
1	27.673	53.614	31	25	66	24	56	22	52	
2	25.601	62.822	28	28	67	28	67	25	64	
3	26.666	62.383	31	26	69	22	58	23	58	
4	25.325	72.053	28	27	73	30	87	25	76	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
494	28.600	70.927	31	24	58	28	71	27	59	
495	23.907	61.134	30	26	62	28	66	24	58	
496	28.229	72.151	31	26	57	28	52	26	55	
497	26.896	62.391	31	29	56	27	56	28	55	

<표 6> 공정 별 예측 적합품률

공정	변수	공정 조정 전		공정 조정 후		예측 적합품률 (%)
		부적합품 수(개)	적합품 예측수(개)	부적합품 예측수(개)	적합품률 (%)	
P_1	C_{11}, C_{12}	497	49	448	9.859	
P_2	C_{21}	497	85	412	17.103	
P_3	C_{31}, C_{32}	497	31	466	6.237	
P_4	C_{41}, C_{42}	497	5	492	1.006	
P_5	C_{51}, C_{52}	497	364	133	73.239	

위 <표 5>의 데이터 #2-1은 P_1 공정의 변수 예측 값과 P_1 공정을 제외한 나머지 공정의 실제 변수 값으로 구성 되어있다. 이와 동일한 방법으로 P_2 공정부터 P_5 공정의 변수 예측 값을 이용하여 공정 별 조정된 데이터 집합(데이터 #2-2~5)들을 구성하였다. 조정된 데이터 집합(데이터 #2-1~5)들을 Model #1을 이용하여 적합품을 예측하고, 예측 적합품률을 비교하여 조정된 공정이 수율 개선에 영향을 주는지 예측하였다. <표 6>는 부적합품 데이터 497개를 공정 별 변수 조정 값으로 조정된 후, Model #1로 적합품을 예측한 결과이다.

여기서, 예측 적합품률이란 신경망을 이용하여 공정을 조정된 후, 부적합품이 적합품으로 바뀔 수 있는 확률을 의미한다. <표 6>의 P_1 공정은 공정 조정 전에 부적합품이 497개였으나, 신경망 예측 모형으로 예측한 변수값으로 P_1 공정 변수를 조정하면 부적합품 497개 중 49개의 제품이 적합품이 될 것으로 예측되었고 예측 적합품률은 9.859%(즉, $49/497 = 0.09859$)로 나타났다. 그리고 다른 공정과 비교하면, P_5 공정이 73.239%로 가장 높

은 예측 적합품률이 나타나고 있다. 만약, P_5 공정을 조정하면, 부적합품 497개중 364개가 적합품으로 바뀔 수 있을 것으로 예측되었다. 본 연구의 예제에서는 예측 적합품률이 가장 높게 나타난 부분 공정을 수율 개선에 영향이 많은 공정으로 선택하였다. 즉, 예측 적합품률이 가장 높게 나타난 P_5 공정인 클리어 공정 내의 온도(C_{51})와 습도(C_{52})를 조정하면 수율 개선에 영향이 많을 것으로 예측하였다. 수율 개선에 영향이 많은 공정을 선택한 후, 공정을 조정하고자 클리어 공정(P_5 공정) 내 온도(C_{51})와 습도(C_{52})의 조정값을 예측하였다. P_5 공정의 예측 조정값은 생산된 제품 1,230개 데이터 중 현재 공정 상태를 반영하고 있는 최근 30개의 데이터를 이용하여 예측하였다. 30개 데이터의 P_1 공정부터 P_4 공정까지 실제 변수인 C_{11} , C_{12} , C_{21} , C_{31} , C_{32} , C_{41} , C_{42} 의 값과 적합품을 의미하는 변수 Y 의 0값을 입력변수로 하여 Model #2~5로 예측하였다. 그 결과, C_{51} 와 C_{52} 의 변수 예측 값 30개를 예측하였다. P_5 공정의 예측 조정 값은 P_1 공정부터 P_4 공정의 다른 조건 때문에 발생한 차이를 고려하여 변수 예측 값들의 평균으로 하였다. C_{51} 의 예측 조정 값은 25.809로, C_{52} 의 예측 조정 값은 66.083로 나타났다. 본 적용 예제에서는 현재 클리어 공정의 온도를 24℃에서 25.809℃로 습도를 58%에서 66.083%로 높여야 수율을 향상시켜 공정의 전체 평균 수율을 유지할 것으로 예측되었다.

3.4 수율 개선 후 확인 단계

수율 개선 단계에서 예측한 정보를 공정에 피드백하여 조정된 후에 수율이 개선되었는지 확인하는 단계이다. 그러나 이 예제에서는 제안한 모형을 실제 현장에 적용하여 수율이 개선되었는지 확인하기가 어렵기 때문에 신경망과 통계적 분석인 다중회귀분석을 이용한 예측을 통하여 이론적으로 수율 개선 여부를 확인하였다. 다중회귀분석의 관계식은 현재 생산된 제품 1,230개 데이터와 분석도구인 미니탭(Minitab)을 이용하여 계산하였다. 다중회귀분석 관계식의 $Adjust-R^2$ 는 67.2%이므로, 관계식을 이용하여 각 제품의 적합여부를 예측할 수 있다. 관계식을 통한 예측값은 0에서 1사이의 값이며 0.5이상이면 부적합품으로 0.5미만이면 적합품으로 판정한다. 예측 정확도는 수집한 공정 데이터 1,230개를 관계식을 이용하여 예측하였으며, 실제 1,230개 데이터 중 943개 데이터를 정확하게 예측하였다. 따라서 예측 정확도는 76.7%로 나타났다.

생산된 제품 1,230개 데이터 중 최근 30개 데이터의 P_5 공정 내 변수들을 제안 모형으로부터 예측된 조정값

으로 조정된 후, 조정된 공정 데이터를 신경망과 통계적 분석인 다중회귀분석으로 적합품을 예측하였다. 신경망을 이용하여 예측하였을 30개 데이터 모두 적합품으로 예측되었고, 다중회귀분석의 관계식 또한 동일하게 30개의 데이터 모두 적합품으로 예측되었다. 공정을 조정된 후 관계식의 예측 정확도가 76.7%를 감안하더라도 제품이 적어도 30개 이상 생산된다면, 생산된 제품이 모두 적합품으로 예상되어 공정 수율이 59.593%에서 60.556%로 개선된다. 만약, 관계식의 예측 정확도가 100%이면 공정을 조정된 후 적합품으로 예측되는 제품이 적어도 20개 이상 생산되면 생산된 제품이 모두 적합품으로 예상되어 공정 수율이 59.593%에서 60.240%로 개선된다.

4. 결 론

실제 자동화 생산라인에서는 수율을 예측하는 능력의 정확도 향상뿐만 아니라 높은 수율을 개선하고 유지시키기 위한 실시간 동적 운용이 가능한 수율 개선, 관리 방법과 모형 개발도 매우 중요하다. 따라서, 본 연구에서는 실시간 및 온라인 기반의 자동화 생산라인에서 신경망을 이용하여 수율을 개선하고 유지가 가능한 동적 수율 개선모형을 제안하였다. 본 연구의 제안 모형은 공정라인에서 수율 저하 시, 신경망 예측 기법을 이용하여 수율 개선에 영향이 많은 공정과 변수 조정값을 예측한 후 공정을 조정하여 공정의 수율을 향상시켜 공정의 전체 평균 수율로 현재 수율을 개선 및 유지가 가능하게 하는 일련의 세부 절차들을 포함한다.

또한, 본 적용 예제에서는 제안 모형의 이해를 돕기 위해 실제 자동차 범퍼 도장 공정의 데이터 적용하였다. 본 예제에서는 제안 모형을 이용하여 수율 개선에 영향을 많이 주는 공정과 변수 조정값을 예측하였다. 그러나 본 제안 모형 내 세부 절차 중 공정 수율 개선 후 확인 단계는 실제 공정에 변수 조정값을 적용해야 하는 현실적 연구 여건의 한계로 인하여 신경망 예측 모형과 통계적 분석인 다중회귀분석의 관계식을 이용한 예측을 통해 이론적으로 그 개선 결과를 확인하였다. 추후 연구로는 실제 현장에 지속적인 공정 피드백을 통해 본 제안 모형을 장기간 적용함과 동시에 다양한 분야로의 모형 적용의 확장을 위한 추가적인 연구를 진행할 예정이다.

참고문헌

- [1] 변성규, 강창욱; “데이터마이닝 기법을 이용한 제

- 조 공정 내의 불량항목별 예측방법”, 한국산업경영 시스템학회지, 27(2) : 10-16, 2004.
- [2] 백동현, 한창희; “데이터마이닝을 이용한 반도체 FAB 공정의 수율개선 및 예측”, 한국지능정보시스템학회 논문지, 9(1) : 157-177, 2003.
- [3] 이장희; “데이터 마이닝 도구의 혼합적용 방법간 수율 예측 성능 비교 연구”, Journal of Business Research, 23(1) : 283-310, 2008.
- [4] Banks, D. L., Parmigiani, G.; “Pre-analysis of super-large industrial data sets,” *Journal of Quality Technology*, 24(3) : 115-129, 1992.
- [5] Guh, R. S., Zorriassatine, F., Tannock, J. D. T., O'Brien, C.; “On-line control chart pattern detection and discrimination-a neural network approach”, *Artificial Intelligence in Engineering*, 13 : 413-425, 1999.
- [6] Kourti, T., Macgregor, J. F.; “Multivariate SPC methods for process and product monitoring,” *Journal of Quality Technology*, 28(4) : 409-428, 1996.
- [7] Mason, R. L., Champ, C. W., Tracy, N. D., Wierda, S. J., Young, J. C.; “Assessment of multivariate process control techniques,” *Journal of Quality Technology*, 29(2) : 140-143, 1997.
- [8] Timothy, J. S., Julie, K. S., Tomas, V.; “The application of artificial neural networks to monitoring and control of an induction hardening process,” *Journal of Industrial Technology*, 16(1) : 1-11, 2000.