

# 데이터 마이닝을 이용한 시멘트 소성공정 질소산화물(NOx)배출 관리 방법에 관한 연구

박철홍\* · 김용수\*\*†

\*경기대학교 일반대학원 산업경영공학과

\*\*경기대학교 산업경영공학과

## A Study on NOx Emission Control Methods in the Cement Firing Process Using Data Mining Techniques

Park, Chul Hong\* · Kim, Yong Soo\*\*†

\*Dept. of Industrial and Management Engineering, Kyonggi University Graduate School

\*\*Dept. of Industrial and Management Engineering, Kyonggi University

### ABSTRACT

**Purpose:** The purpose of this study was to investigate the relationship between kiln processing parameters and NOx emissions that occur in the sintering and calcination steps of the cement manufacturing process and to derive the main factors responsible for producing emissions outside emission limit criteria, as determined by category models and classification rules, using data mining techniques. The results from this study are expected to be useful as guidelines for NOx emission control standards.

**Methods:** Data were collected from Precalciner Kiln No.3 used in one of the domestic cement plants in Korea. Thirty-four independent variables affecting NOx generation and dependent variables that exceeded or were below the NOx emission limit (>1 and <0, respectively) were examined during kiln processing. These data were used to construct a detection model of NOx emission, in which emissions exceeded or were below the set limits. The model was validated using SPSS MODELER 18.0, artificial neural network, decision tree (C5.0), and logistic regression analysis data mining techniques.

**Results:** The decision tree (C5.0) algorithm best represented NOx emission behavior and was used to identify 10 processing variables that resulted in NOx emissions outside limit criteria.

**Conclusion:** The results of this study indicate that the decision tree (C5.0) can be applied for real-time monitoring and management of NOx emissions during the cement firing process to satisfy NOx emission control standards and to provide for a more eco-friendly cement product.

**Key Words:** Firing process, NOx, Decision tree, Classification matrix table, ROC curve

• Received 31 August 2018, 1st revised 6 September, accepted 7 September 2018

† Corresponding Author(kimys@kgu.ac.kr)

© 2018, The Korean Society for Quality Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-Commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

※본 연구는 경기도의 경기도 지역협력연구센터 사업의 일환으로 수행하였음[GRRCK2018-B05, 스마트제조 응용 기술 연구]

## 1. 연구목표 및 배경

우리나라의 경우 시멘트산업은 철강·석유화학산업 등과 함께 대표적인 국가 기간산업의 하나로 건축·토목산업에 있어서 필수 불가결한 기초소재산업이며, 짧은 기간 동안에 생산기술과 설비가 세계 정상급으로 발전하여 왔다. 우리나라는 세계 11대 시멘트 생산국이고 소비 기준으로 세계 10대 시멘트 소비국이다(International Cement Review Magazine, 2011).

최근에는 수입 시멘트의 증가, 슬래그 등 혼화재의 사용비중 확대, 스틸하우스 등 대체재의 시멘트 수요 잠식 등 안팎으로 어려움을 겪고 있으며 점차 환경에 대한 일반의 관심이 증가하면서 친환경 산업으로의 역할 강화가 요구되는 등 향후 시멘트 산업의 활로를 새롭게 모색해야 하는 중요한 시기로 시멘트 산업의 경쟁력 구축이 절실히 요구되는 상황이다.

최근 환경부는 질소산화물(NOx, 이하 Nox)에 kg당 2130원의 배출 부담금을 물리는 ‘대기환경보전법 시행령 개정안’을 입법 예고했다. 2019. 1월 부터 NOx 배출 허용 기준이 종전 330ppm에서 270ppm으로 강화됨에 따라 개정안대로 시행되면 국내 A시멘트 기업의 경우 연간 약 100억원에 달하는 부과금이 예상된다.

시멘트 산업은 생산 및 품질을 유지하기 위해서는 고온 소성이 필수적이기 때문에 불가결하게 NOx가 발생하는 조건을 가지고 있다. 국내 A시멘트 기업은 이미 질소산화물 최적 방지시설인 선택적비촉매환원설비(SNCR: Selective Non-catalytic Reduction)를 설치·운영하고 있지만 기술적으로 추가 저감에 현실적으로 어려움이 있다.

최근 이슈화 되고 있는 4차 산업혁명 시대에 국내 시멘트 제조 기업들은 설비, 생산, 품질, 에너지, 환경 등 전 부문에서 어떻게 기존 공장을 스마트 공장으로 변환할 것인가에 대한 구축 방법과 실행방안이 요구되어지고 있다.

따라서 본 연구에서는 스마트팩토리 실현 방안의 일환으로 시멘트 소성공정의 NOx 배출 허용기준 초과·미만 분류모형, NOx 배출량 초과 규칙(rule) 및 NOx 발생에 영향을 미치는 주요 변수를 찾고자 Preheater Kiln 형태인 Kiln3호기를 대상으로 데이터를 수집하였다. NOx 발생과 관련이 있는 소성공정 변수들과 7개 대기오염물질을 24시간 측정할 수 있는 굴뚝 원격 감시(TMS: Tele-Metering System) 정책에 따라 설치된 TMS에서 측정된 NOx 데이터를 수집하였다.

수집된 데이터를 사용하여 데이터 마이닝 기법인 Decision Tree(의사결정나무 C5.0), ANN(인공신경망), Logistic regression(로지스틱 회귀분석)를 이용하여 분류모형의 유용성 검증을 거쳐 NOx 배출 허용기준 초과·미만 분류모형, NOx 배출량 초과 규칙(rule) 및 NOx 발생에 영향을 미치는 주요 변수를 제시하고자 한다.

본 연구 결과를 통해서 시멘트 소성공정 Kiln 3호기에서 실시간으로 NOx 배출량 허용기준 만족여부 모니터링 및 NOx 배출량을 관리하기 위한 기준으로 활용하고자 한다

## 2. 관련 문헌연구

최근 전세계적으로 환경규제의 강화로 대기오염 물질의 주요 원인인 연소장치에서의 오염 물질 배출을 줄이기 위한 많은 노력이 행해지고 있다. 연소 가스 중 대표적인 오염 물질으로 일산화탄소(CO), 황산화물(SOx), 질소산화물(NOx) 등이 있다. 특히 NOx는 산성비와 광학 스모그의 원인이고 인체와 동식물에 심각한 영향을 미칠 수 있다(Kang et al., 2007).

시멘트산업에서 배출되는 CO의 환경영향 예비실태조사 결과 내륙지역에 위치한 시멘트 사업장이 위치한 주변지역의 이산화질소 농도분포가 집중되어 나타나고 있고, 해안 지역 시멘트 사업장을 중심으로 한 주변지역 이산화 질

소 농도 분포가 사업장 중심 주변지역이 타 지역에 비해 높은 분포를 나타내었다(Kim et al., 2003).

시멘트 제조공정에서 NOx는 고온에서 생성되는 NOx(Thermal NOx), 연료로부터 발생하는 NOx(Fuel NOx), 화염면 및 근방에서 발생하는 NOx(Prompt NOx), 원료로부터 발생하는 NOx(Feed NOx)로 생성된다. 대부분의 경우 Thermal NOx, Fuel NOx가 중요하다.

시멘트 제조공정은 1450°C의 고온 공정인 소성공정을 포함하고 있다. 이런 고온의 공정은 공기중의 질소와 연료 중의 N를 산화시켜 NOx가 발생됨에 따라 저감 방법으로 다단연소는 하소로를 가지는 NSP Kiln 시스템 적용시 40% 정도의 NOx 저감효과가 있는 것으로 나타났다(Lim et al., 1998).

시멘트 소성공정의 NOx 저감을 위해 배가스 온도가 1000°C 전후인 경우 유리한 선택적비촉매환원(SNCR: Selective Noncatalytic Reduction) 시스템을 적용하여 환원제 분사조건의 최적화를 검토하였다. 그 결과로 환원제 분사지점별 분사량에 따른 NOx 제거효율 실험결과 분사지점 부근을 통과하는 배가스의 유동특성에 따른 환원제와 배가스의 혼합 정도에 의존하는 것으로 확인 되었다. 환원제 분사노즐 형태에 따라서는 액적 크기는 50~100 $\mu$ m 원뿔형 노즐(full cone nozzle) 이 부채형 노즐(flat spray nozzle)보다 NOx 제거 효율이 우수하였다. 요소수용액 농도에 따른 NOx 제거효율 실험결과 환원제 수용액의 농도가 낮을수록 NOx 제거에 유리하였다(Ham, 2007).

시멘트 Precalciner kiln에서 NOx 배출량을 예측하고 최적화하기 위해 데이터 마이닝 기법인 ANN(인공신경망)을 사용하여 NOx 예측 모델을 수립하였고 모델의 정확도는 MRE가 2% 미만이고, 실제값과 예측값에 대한 상관계수(r)는 0.9290로 정확도는 우수하였다. 유전자 알고리즘(genetic algorithm, GA)을 사용하여 NOx 배출량을 최적화하였다. 민감도(sensitivity) 분석결과 소성로 온도(furnace temperate), 원료량(raw meterial quantity) 및 3 차 공기 온도(third air temperate)가 질소 산화물 배출 농도에 더 큰 영향을 미치는 주요 운전변수로 나타났다. 데이터마이닝 기법 ANN, GA를 사용하여 NOx 배출량을 효과적으로 예측하고 최적화 할 수 있었다(Zhang, 2017).

건식 시멘트 Kiln에서 석탄 또는 코크스 대신 소성로에서 대체 연료로 폐기물 가공 원료(RDF: Refuse Derived Fuel)를 사용하면 화염온도 및 공기 과잉이 낮고 화석연료 대비 질소 함량이 낮기 때문에 NOx 발생량을 줄일 수 있다. 하지만 폐기물 연소는 직접 연소, 가스화 연소(gasification combustion), 복합적 화학 반응 속도론(complex chemical kinetics), 점화 제어 등 몇 가지 기술적 과제에 직면한다(Liu et al., 2014).

시멘트 소성공정에서 부원료로 하수 슬러지를 사용한 결과 하수 슬러지에서 배출 된 NH3 양에 의존하는 NOx 배출을 줄이는 데 도움을 주었다. 또한 배가스 내의 NOx와 NH3 농도는 음의 상관 관계가 존재하여 선택적비촉매환원(SNCR)시스템에서 암모니아수의 주입량이 증가함에 따라 NOx 제거 기여도는 감소한다. 따라서 시멘트 Kiln에서 하수 슬러지를 공동 처리하는 동안 SNCR 시스템에서 암모니아수의 주입량이 감소하여 운영비를 절감 할 수 있다고 제안하였다(Lv et al., 2016).

데이터마이닝이란 대용량의 데이터로부터 유용하게 활용될 수 있는 지식을 효과적으로 찾아내는 지식 탐사의 한 연구 분야로서 데이터로부터 정보를 추출하기 위해 여러 가지 기법을 적용하는 특정단계로 정의할 수 있다(Song et al., 2006). 제조 공정에서의 마이닝의 적용은 CRM에서의 마이닝보다 현실적으로 많은 어려움이 존재하기 때문에 그다지 활성화되어있지 않은 것이 사실이나, 무결점의 최상의 제품을 만들기 위해서는 반드시 필요하다(Lee and Nam, 2006). 데이터 마이닝 기법 중에서도 최근에 각광을 받는 분야는 인공신경망(ANN)으로서 패턴 인식이나 예측 분야에 많이 응용되고 있다(Lee et al., 2001). 시멘트 공정에서는 Maiko OHNO 소성공정에 대해 인공신경망(ANN)을 이용한 시멘트 품질 예측 시스템 개발 연구가 있었고, 그 결과는 이전의 연구와 일반적으로 일치하고 정확도는 기존의 방법에 비해 향상되었다(Ohno, 2012).

NOx 배출 저감 기술은 주로 장비 개조, 촉매 환원 및 운전변수 제어 방법들에 대한 연구 결과들이 있다. 그러나 국내 시멘트 업종에서 데이터마이닝 기법을 활용한 NOx 배출 허용기준 초과·미만 분류모형, NOx 배출량 초과 규

칙(rule) 및 NOx 발생에 영향을 미치는 주요 변수를 찾고자 하는 연구는 많이 진행되지 않은 것으로 판단된다. 따라서, 본 연구에서는 NOx 발생에 영향을 미치는 주요 변수를 데이터 마이닝을 통해 찾고 그 의미를 분석하고자 한다.

### 3. 시멘트 소성 공정 소개

본 연구는 국내 A시멘트 기업에서 Nox를 발생시키는 소성공정의 Preheater Kiln 형태인 Kiln 3호기를 대상으로 하였다. 소성공정은 Figure 1과 같은 과정을 거친다.

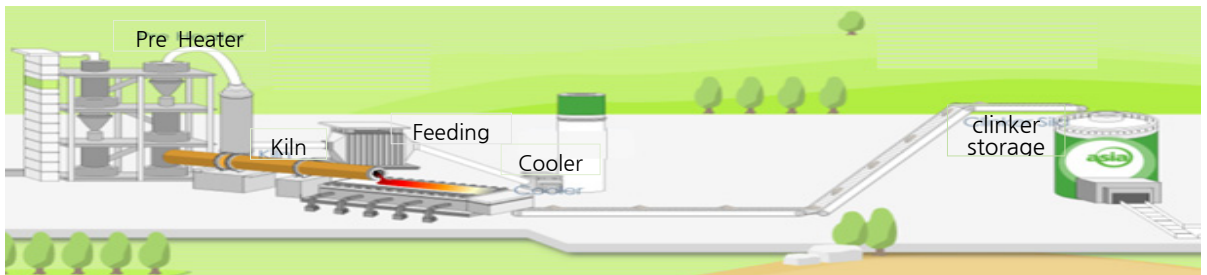


Figure 1. Cement kiln process

원료공정에서 제조한 조합원료는 예열기(Preheater)를 통하여 소성로에 공급된다. 공급된 원료는 미분탄 등의 연료를 사용하는 대형 버너에 의해 소성로에서 1450℃의 온도에서 소성되어 클링커(clinker)라는 새로운 광물이 생성되어 진다. 소성로(Kiln)는 내부에 내화연와를 쌓은 강철제의 커다란 원통으로 3~5도의 경사를 두어 1분에 3~4회의 속도로 회전한다. 생성된 클링커는 냉각장치(Cooler)로 이송되어 공기에 의해 급냉 시킨다. 냉각된 클링커는 클링커 저장장(Clinker storage)된다. 이런 과정에서 시멘트 제조 공정의 특성상 고온 소성 과정에서 질소가 산화되어 NOx가 발생한다.

A사의 경우 소성공정의 Pre heater, Kiln, Cooler subprocess에서 온도, 압력 등의 값을 MES에 초단위로 수집하고 TMS에서 측정된 NOx는 CMS에 5분 단위로 수집 모니터링하고 있으나 이에 대한 분석은 미흡한 상황이다.

따라서 본 연구에서는 MES에서 NOx 발생과 관련있는 Kiln 3호기 공정변수들과 CMS에서 NOx 데이터를 수집하였다. NOx 배출 허용기준 초과·미만 분류 모형 개발, NOx 발생 규칙(Rule) 및 NOx에 영향을 미치는 주요 변수를 찾아 보고자 한다. 이 결과를 이용 NOx 배출의 지속적인 모니터링 및 관리기준으로 활용하고자 한다.

## 4. 연구방법 및 절차

### 4.1 연구모형

본 연구에서는 국내 A시멘트 기업의 친환경 제품 생산을 하기 위해 소성공정 Kiln 3호기 공정변수와 NOx 배출량 데이터를 이용하였다. 데이터 마이닝 기법을 이용하여 제조과정에서 발생하는 NOx 배출량에 대한 허용 기준 초과·미만 분류모형을 개발하고자 한다. 모델간의 분류 모형 성능평가를 하여 가장 우수한 모델을 선정하고자 한다.

본 연구의 분석 프로세스는 시멘트 및 소성공정의 이해, 데이터 수집, 데이터 전처리, 모델링, 소성공정 설비고장 발생 검출 모형 평가의 단계로 구분되며 Figure 2와 같다.

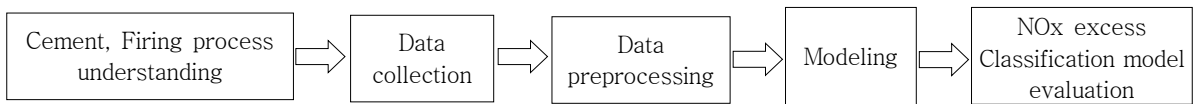


Figure 2. Research flow chart

데이터 수집 단계에서는 국내 A시멘트 기업의 NOx와 관련이 있는 소성공정(Preheater, Kiln, Cooler)의 공정변수들과 NOx 배출량에 대해 최근 2개월(18.4~6월) 동안의 데이터를 추출 하였다.

데이터 전처리 단계에서는 데이터에 대한 이해를 통해 설비 비가동 및 설비 보수시 데이터인 이상치는 제거하고, 결측치 처리 등을 실행하였다. 모델링 단계에서는 SPSS Modeler 18.0로 노드를 구성하여 모델을 수립 하였다. 분류 모형 평가단계에서는 의사결정나무(C 5.0), ANN(인공신경망), 로지스틱 회귀분석 기법에서 도출된 모형을 Test set data에 적용하여 모델의 예측력은 분류행렬표 및 ROC 곡선으로 비교 평가 하였다.

### 4.2 데이터 수집

데이터 수집은 국내 A시멘트 기업의 Kiln 3호기 NOx 배출 허용기준 초과·미만 분류 모형을 생성하기 위해 데이터를 추출하는 단계이다. 본 연구에서는 소성공정의 Preheater, Kiln, Cooler subprocess에서 NOx와 관련 있는 공정변수 데이터와 NOx 데이터를 5분 단위로 수집하였다.

Table 1과 같이 Pre heater 공정에서 20개 변수, Kiln 공정에서 5개 변수, Cooler 9개 변수 총 34개의 독립변수와 NOx 배출 허용 기준 초과·미만(초과1, 미만0) 종속변수를 대상으로 분석하였다.

Table 1. Processing variables

process	Variable description	Var.type	process	Variable description	Var.type
Pre-heater	X1. Raw material	Input amout(t/hr)	Pre-heater	X18. Cyclone Pressure	Pressure
	X2. KO2 Rear pressure	Pressure		X19. KO2 Damper	%
	X3. Pre'coal	Input amout(t/hr)		X20. KO2 kilowatt	kW
	X4. Inlet by-product	Input amout(t/hr)	Kiln	X21. Primary Fan Pressure	Pressure
	X5. Softplastic(Riser)	Input amout(t/hr)		X22. Focus temperature	Temperature (°C)
	X6. Softplastic(MFC)	Input amout(t/hr)		X23. Kiln Amp	Ampere
	X7. Renewable fuel oil	Input amout		X24. Hood Pressure	Pressure
	X8. Sewage sludge	Input amout	Cooler	X25. Main Coal	Input amout
	X9. T/A Duct	Temperature(°C)		X26 Grate Drive	spm
	X10. T/A Pressure	Pressure		X27. Grate Drive cooler1	%
	X11. MFC Damper	%		X28. Grate Drive cooler2	%
	X12. MFC Bed Pressure	Pressure		X29. Grate Drive cooler3	%
	X13. MFC Outlet temperature	Temperature(°C)		X30. Grate Drive cooler4	%
	X14. 4 stages Inlet line	Temperature(°C)		X31. C/Fan Pressure1	Pressure
	X15. 4stagesMFCline	Temperature(°C)		X32. C/Fan Pressure2	Pressure
	X16. KO2 Temperature	Temperature(°C)		X33. C/FanPressure3	Pressure
	X17. Riser Temperature	Temperature(°C)		X34. C/Fan Pressure4	Pressure
Y	NOx emission standard	Binary[Under(0), Excess(1)]			

### 4.3 데이터 전처리

데이터 전처리는 데이터마이닝 분석 목적에 따라 데이터를 처리하는 일련의 과정을 말한다. 실제 데이터는 불완전(incomplete), 잡음(noisy), 불일치(inconsistent)하기 때문에 데이터의 전처리가 필요하다. 일반적으로 데이터 전처리 단계에서는 데이터 중 결측치를 채워넣고, 잡음이 있는 데이터를 제거하며 이상치를 식별하고, 데이터 불일치를 교정한다.

실제 데이터베이스는 일반적으로 대용량이며 여러 개의 이질적인 데이터 원천에서 취합되기 때문에 노이즈에 취약하며 데이터의 일관성이 떨어진다. 부정확, 불완전, 비일치성 데이터는 대량의 현실 데이터 베이스와 데이터웨어하우스의 일반적인 특성이다.

따라서 데이터 전처리가 필요하며 몇 가지 방법으로는 데이터 정제(데이터 전처리에는 데이터내의 노이즈를 제거하고 일관성 결여를 교정하는데 사용), 데이터 통합(여러 데이터 원천에서 발생하는 데이터를 일관성 있는 데이터 저장소인 데이터웨어하우스에 저장하는 과정), 데이터 축소(중복제거, 군집화와 같은 과정을 통하여 데이터의 크기를 축소하는 과정), 데이터 변환(정규화)의 방법들이 있다.

본 연구에서 데이터 정제하는 작업으로 결측치 및 NOx 100ppm 이하 비정상 데이터가 있는 경우 해당 레코드를 삭제하고 설비 비가동, 보수시 발생한 이상치 데이터는 제거하여 일관성이 없는 문제점을 최대한 해결하였다. 전처리 결과 독립변수 34개, 종속변수 1개에 대한 데이터 레코드는 12,680개이다.

### 4.4 모델링

본 연구에서는 모델링은 데이터 마이닝 도구인 SPSS MODELER 18.0을 사용하여, 데이터마이닝 수행노드는 Figure 3과 같다. 본 연구에 사용된 모델링 및 노드에 대한 설명은 다음과 같다.

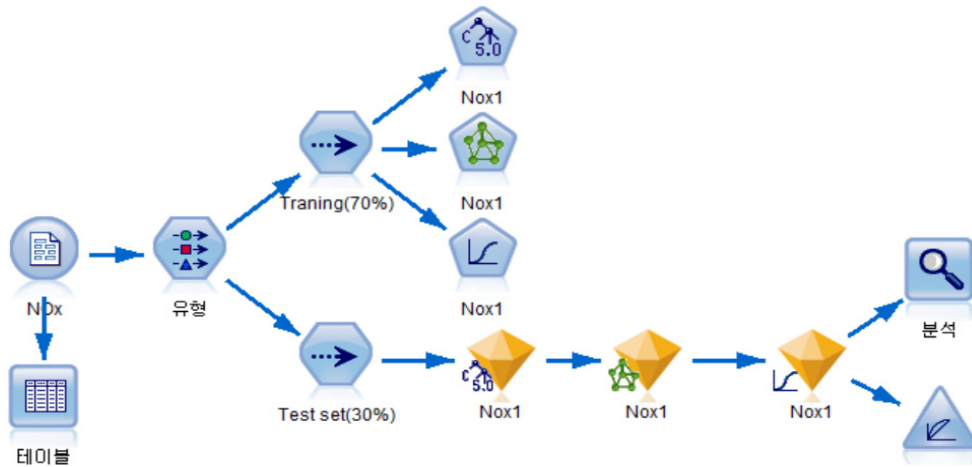


Figure 3. Data mining perform node

입력노드는 국내 A시멘트 기업 시멘트 제조시 질소산화물 발생과 상관이 있는 소성공정의 공정변수와 NOx 초과·미만(초과1, 미만0)을 대상으로 분석 하였다.

유형노드는 데이터 유형을 지정, 변경할 수 있는 노드로 독립변수는 연속형, 종속변수는 플레그로 지정하였고 Y에 대해서는 목표로 지정하였다. 데이터 레코드는 12,680개로 훈련 집합(training data set)과 테스트 집합(test data set)을 70%, 30%로 분할하였다.

C5.0 노드는 의사결정나무 규칙을 생성하기 위해 정보이득률(gain ratio)을 계산하여 가장 큰 비율값을 가진 변수를 선택하여 데이터를 분리하여 나무모델을 만든다. 신경망 노드는 신경망 학습을 통해 생성된 속성을 내재하고 있는 모델로 테스트 집합에 대해 이 노드를 지나게 되면 생성된 모형에 따라 예측값을 생성한다. 로지스틱 회귀분석 노드는 2개 이상의 이산형 값을 갖는 종속변수와 독립변수들 간의 인과 관계를 로지스틱 함수를 이용하여 추정하는 모델링 노드이다(Kim, 2011).

분석노드는 모델의 분류성능을 평가하는데 사용되며 정확성 척도는 분류행렬(classification matrix)로부터 유도된다. 예측 값이 실제 목표값을 얼마나 정확하게 예측하는지 비교하는데 사용된다. 평가모드는 ROC곡선(receiver operating characteristic), 이익도표 등 예측모형의 정확성을 평가하는 그래프로 사용된다. 본 연구에서는 분류행렬 표 및 ROC곡선을 통해 모델의 예측력을 평가 하였다.

### 4.5 모델평가

본 연구에서는 모형의 정확성 척도를 산정하고자 분류행렬 및 ROC곡선을 사용하였다. 분류행렬은 분류기가 특정 데이터 세트에 대해서 산출하는 정확한 분류와 부정확한 분류를 요약한다. 분류 오류의 추정값을 얻기 위해서 테스트 집합(Test data set)으로부터 계산된 분류행렬을 사용하였다.

분류행렬의 행과 열은 각각 실제 클래스와 예측된 클래스에 해당한다. 본 연구에서는 C<sub>0</sub>는 NOx 배출 허용기준 미만, C<sub>1</sub>은 NOx 배출 허용기준 초과로 표시하였다. Table 2는 분류행렬표를 나타낸 것이다.

본 연구에서는 정확도(Accuracy), 에러율(error rate), 민감도(sensitivity), 특이도(specificity)로 분류모형의 유용성을 검증하였다. 분류행렬표의 대표적인 측정치는 다음과 같다.

- (1) 정확도(Accuracy)는 전체집단에서 각각의 집단을 정확하게 분류하는 정도를 나타내고 (TP+ TN)/P+N으로 측정된다.
- (2) 에러율(error rate)는 전체집단에서 각각의 집단을 잘못 분류하는 정도를 나타내고 (FP+ FN)/P+N으로 측정된다.
- (3) 민감도(sensitivity)는 긍정 집단을 얼마나 정확히 인식했는지 알려주는 지표로 실제 C<sub>0</sub> 집단을 C<sub>0</sub> 집단으로 정확하게 분류할 확률로써 TP/P로 측정된다.
- (4) 특이도(specificity)는 부정 집단을 얼마나 정확히 인식했는지 알려주는 지표로 실제 C<sub>1</sub> 집단을 C<sub>1</sub> 집단으로 정확하게 분류할 확률로써 TN/N로 측정된다.
- (5) 위양성률(false positive rate)은 C<sub>0</sub>집단으로 분류된 레코드 중에서 실제 C<sub>1</sub> 집단을 C<sub>0</sub> 집단으로 잘못 분류한 레코드의 비율을 의미하며, FP/P'로 측정된다.
- (6) 위음성률(false negative rate)은 C<sub>1</sub> 집단으로 분류된 레코드 중에서 실제 C<sub>0</sub>집단을 C<sub>1</sub> 집단으로 잘못 분류한 레코드의 비율을 의미하며, FN/N'로 측정된다.

Table 2. Classification matrix table

		Predicted class		Total
		C <sub>0</sub> (Under)	C <sub>1</sub> (Excess)	
Actual class	C <sub>0</sub> (Under)	TP(True Positive) Number of correctly categorized C <sub>0</sub> data	FN(False Negative) Number of C <sub>0</sub> data misclassified as C <sub>1</sub>	P
	C <sub>1</sub> (Excess)	FP(False Positive) Number of C <sub>1</sub> data incorrectly classified as C <sub>0</sub>	TN(True Negative) Number of C <sub>1</sub> data correctly classified	N
Total		P'	N'	P+N



## 5. 분류모형 수립 및 평가

### 5.1 신경망, 의사결정나무, 로지스틱 회귀분석 분석 결과

본 연구를 위해 데이터마이닝 도구인 SPSS Modeler 18.0을 사용하였으며 의사결정나무(C5.0), 신경망, 로지스틱 회귀분석 모형에 Test set data에 적용 결과 분류 행렬표, 정확도, 에러율, 민감도, 특이도 비교 결과는 Table 3, Table 4와 같다.

NOx 배출 허용기준 초과·미만에 대해 의사결정나무(C5.0) 분류모형 적용 결과 정확도 94.22%, 에러율 5.78%, 민감도 88.86%, 특이도 97.09%로 3가지 기법 중 가장 우수하게 나타났다. 신경망 분류모형 적용 결과 정확도 89.97%, 에러율 10.03%, 민감도 83.00%, 특이도 93.71%로 3가지 기법 중 두 번째로 우수하게 나타났다. 로지스틱 회귀분석 분류모형 적용 결과 정확도 86.60%, 에러율 13.40%, 민감도 77.06%, 특이도 91.72%로 다른 결과 값에 비해 낮음을 알 수 있다.

의사결정나무는 93% 이상의 높은 정확도와 민감도를 나타난 반면에 신경망과 로지스틱 회귀분석은 정확도와 에러율에서 차이를 보이고 있다. 특이도는 세 가지 기법 모두 90% 이상으로 나타 났지만 의사결정나무가 가장 높은 결과 값을 보이고 있다.

**Table 3.** C5.0, Neural network, Logistic regression classification matrix table

Decision tree(C5.0)		Predicted class		Total
		0(under)	1(excess)	
Actual class	0(under)	1,197	150	1,347
	1(excess)	73	2,439	2,512
Total		1,270	2,589	3,859
Accuracy=94.22%, error rate=5.78%, sensitivity=88.86%, specificity=97.09%				

Neural network		Predicted class		Total
		0(under)	1(excess)	
Actual class	0(under)	1,118	229	1,347
	1(excess)	158	2,354	2,512
Total		1,306	2,659	3,859
Accuracy=89.97%, error rate=10.03%, sensitivity=83.00%, specificity=93.71%				

Logistic regression		Predicted class		Total
		0(under)	1(excess)	
Actual class	0(under)	1,038	309	1,347
	1(excess)	208	2,304	2,512
Total		1,246	2,613	3,859

Accuracy=86.60%, error rate=13.40%, sensitivity=77.06%, specificity=91.72%

**Table 4.** Comparison of classification matrix table results of each technique

	Decision tree(C5.0)	Neural network	Logistic regression
Accuracy	94.22%	89.97%	86.60%
error rate	5.78%	10.03%	13.40%
sensitivity	88.86%	83.00%	77.06%
specificity	97.09%	93.71%	91.72%

의사결정나무는 94% 이상의 높은 정확도를 나타난 반면에 신경망과 로지스틱 회귀분석은 정확도와 에러율에서 차이를 보이고 있다. 특이도는 세 가지 기법 모두 91% 이상으로 나타났으나 의사결정나무가 가장 높은 결과 값을 보이고 있다.

모형을 평가할 때에는 위 네가지 평가 측도를 고려한 평가가 이루어져야 한다. 따라서 의사결정나무(C5.0) 모형이 정확도, 에러율, 민감도에서도 가장 우수한 것으로 평가되었다.

일반적으로 민감도와 특이도를 동시에 증가시키는 것은 불가능하다. 민감도를 높이면 특이도가 감소하고 또한 반대가 성립하게 된다. 민감도와 특이도를 동시에 나눈 것 중 ROC(receive operating characteristic) 곡선이 널리 사용되는데 이는 분류기의 경계치를 조정하여 가면서 (1-특이도)(false positive rate)를 x축에 민감도를 y축에 도식한 것이다. 적합한 모형이 종속변수의 값이 '1'인 개체와 '0'인 개체를 얼마나 잘 식별할 수 있는지를 재는 측도라고 할 수 있다.

의사결정나무 모형, 신경망 모형, 로지스틱 회귀 모형의 ROC 곡선은 Figure 4와 같다. 가장 아래의 대각선은 민감도와 1-특이도(위 양성률)이 정확하게 일치하는 지점을 보여준다. ROC 곡선이 대각선에 가까울수록 모델의 정확도가 떨어진다는 뜻이다. 따라서 의사결정나무 모형이 가장 정확하게 나타났다.

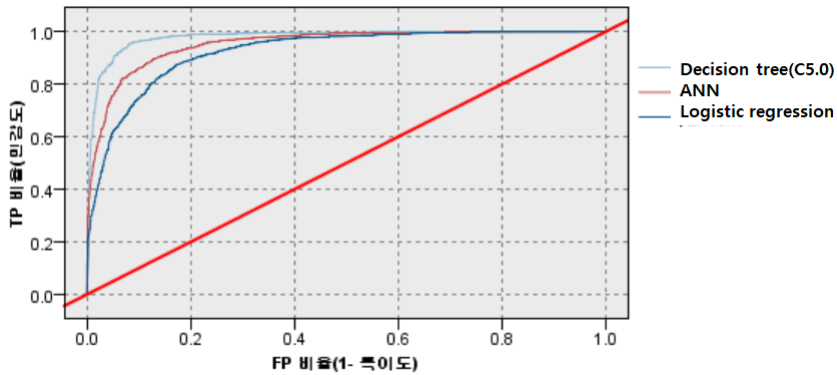


Figure 4. ROC Curve of Decision Tree(C5.0), ANN, Logistic regression

## 5.2 NOx(질소산화물) 분석결과

본 연구에서는 시멘트 소성시 발생하는 NOx(질소산화물) 배출 허용기준 초과·미만을 분류하기 위해 검출 모형을 생성하였다. 소성공정은 Feeding(Raw Mill 반제품 투입 공정), Preheater(Raw Mill에서 생산된 로믹스를 예열하는 설비(1~5단: 300~900도), Kiln, CoalLine(킬런 소성온도관리를 위해 원료 투입), Cooler, IDFan(Kiln에서 발생한 Hot gas를 배출하는 역할)로 구성되어 있다. 각 하위공정에서 생성되는 온도, 압력 등을 독립변수로 선정하고 NOx 배출량 초과(1), 미달(0)을 종속변수로 지정하였다.

의사결정나무 분석 모형에서 소성 공정에서 NOx 기준 배출량 초과 분류 규칙(rule)은 배출 분류 규칙 10개는 Figure 5와 같다. 첫 번째 분류 규칙을 보면 X20(KO2전력)이 2,001 이하이고, X19(KO2 Damper)가 52이하일 때와 X20(KO2전력)이 2,001 초과하고, X24(Hood Pressure)가 -1이하이고, X30(Grate Drive cooler4)이 60이하 일 때 NOx 배출 허용기준을 초과함을 알 수 있다. 의사결정나무(C5.0) 분석결과 규칙1의 4단계까지의 의사결정나무 구조는 Figure 6과 같다.

이런 결과로 NOx 배출 초과에 영향을 미치는 주요 변수는 X20(KO2 전력), X30(Grate Drive cooler), X22(Focus temperature), X23(Kiln Amp), X1(Raw material amount), X10(T/A Pressure), X17(Riser Temperature), X21(Primary Fan Pressure), X16(KO2 Temperature), X12(MFC Bed Pressure)로 나타났다.

본 연구 결과를 통해 제안된 의사결정나무(C5.0) 모형을 적용하여 NOx 배출 모니터링 및 관리 기준으로 활용할 수 있을 것으로 기대되며 다른 Kiln 호기에 대해서도 추가 연구가 요구된다.



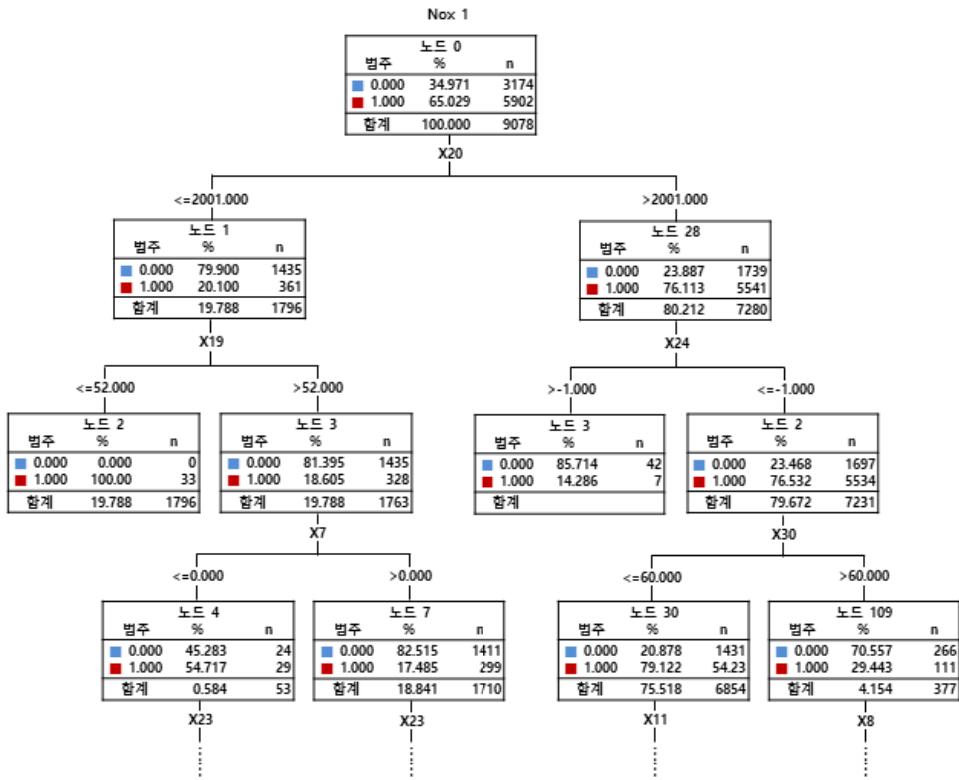


Figure 5. C5.0 decision tree of rule 1

## 6. 결론 및 추후 연구과제

최근 환경부가 2017. 9월 발표한 ‘미세먼지 관리 종합대책’의 후속조치로 질소산화물 등 미세먼지 배출 규제를 강화하는 ‘대기환경보전법 시행규칙 일부 개정령안’을 입법 예고함에 따라 대기환경오염물질 배출 사업장의 질소산화물(NOx) 배출허용기준이 강화된다. 이번 개정안은 2019년 1월 1일부터 적용될 예정이다.

시멘트 제조업의 경우 NOx 배출량 허용기준이 기존 330ppm에서 270ppm으로 NOx(질소산화물) 배출규제가 강화된다. 따라서 국내 A시멘트 기업의 경우 여러 방법으로 NOx(질소산화물) 배출 감소를 위한 개선활동을 추진하고 있지만 가시적인 성과가 나타나지 않고 있다. 시멘트 업계는 대부분 공정상에서 많은 데이터들을 수집하고 있으나 데이터마이닝에 대한 이해 부족으로 이를 활용하지 못하고 있는 것이 현실이다.

따라서 본 연구는 국내 A시멘트 기업의 소성 Kiln 3호기 공정변수와 NOx 배출량에 대해 수집된 데이터를 이용하여 데이터 마이닝 기법인 의사결정나무(C5.0), 인공신경망, 로지스틱 회귀분석을 이용하였다.

이 결과인 분류 행렬표에서 정확도, 에러율, 민감도, 특이도 및 ROC곡선을 통해 모델의 예측력을 분석한 결과 의사결정나무(C5.0), 인공신경망, 로지스틱 회귀분석 기법들 중 의사결정나무(C5.0) 모형의 결과 값이 가장 우수한 성능을 나타내었다. 이 결과로 NOx 배출 허용기준 초과·미만 분류 모형, NOx 초과 분류 규칙(rule) 10개 및 NOx 배출에 영향을 미치는 주요변수 10개를 제안하였다.

본 연구 결과를 통해 제안된 의사결정 나무 분류 모형으로 NOx 배출 관리를 위한 기준으로 유용할 것으로 기대된

다. 최근 NOx 배출 표준을 만족시키는 경제적인 방법으로 연소의 운전 변수를 최적화 하는 것이 국제적인 관심을 끌고 있다(Zhang, 2017). 이런 상황에서 대기오염물질인 NOx 발생을 감소하기 위해 데이터마이닝과 DOE를 이용한 공정변수 최적화에 대한 추가 연구가 병행되어야 할 것이다.

## REFERENCES

- Ham, S. W. 2007. "Application of SNCR Process for NOx Removal from Limestone Calcining Process." *Research Review Kyungil University* 16(6):995-1002.
- International Cement Review Magazine. 2011. "The Global Cement Reort 9th Edition." *International Cement Review*
- Kang, S.-K., Azimov, U. B. and Kim, K.-S. 2007. "NOx Reduction Study in Oscillating Combustion Burner." *Journal of Korean Society of Combustion* 12(4):22-30.
- Kim, J. K., et al. 2011. "A Study on Quality Control Using Data Mining in Steel Continuous Casting Process." *Journal of the Korea Society of IT Services* 10(3):113-126.
- Kim, Y. S., et al. 2003. "PA17) A pilot study of NOx concentration emitted from cement industry." *Proceeding of the 35th Meeting of KOSAE* 235-236.
- Lee, H. and Nam. H. S. 2006. "A Quality Data Mining System in TFT-LCD Industry." *Journal of the Korean Society for Quality Management* 34(1):13-21.
- Lee, J. H., Yu, S. J. and Park, S. C. 2001. "Design of Intelligent Material Quality Control System based on Pattern Analysis using Artificial Neural Network." *Journal of the Korean Society for Quality Management* 29(4):38-53.
- Lim, K. K., et al. 1998. "Process Analysis for DeNOx in Cement Calcination Process." *Proceeding of the Meeting of KOSAE* 146-147.
- Liu, H. B., Chen, X. D. and Gu, J. 2014. "NOx reduction review on fuel alternative in cement kiln." *Advanced Materials Research* 864:1626-1629.
- Lv, D., et al. 2016. "Effects of co-processing sewage sludge in cement kiln on NOx, NH3 and PAHs emissions." *Chemosphere* 159: 595-601.
- Ohno, M., Kurokawa, D. and Hirao, H. 2012. "Establishment of a cement quality predictive system-Analyses of the relationships between characteristics ana properties of cement-." *Cement Science and Concrete Technology* 66(1):87-94.
- Song, S.-I., Jung, H.-J. and Shin. S.-M. 2006. "Optimization Methodology Integrated Data Mining and Statistical Method." *Journal of the Korean Society for Quality Management* 34(4):33-41.
- Zhang, Y., et al. 2017. "ANN-GA approach for predictive modelling and optimization of NOx emissions in a cement precalcining kiln." *International Journal of Environmental Studies* 74(2):253-261.