



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학전문석사학위논문

석탄화력 발전소
탈황공정 운전 Data를 활용한
성능예측모델 적용 연구

**Prediction Model of Desulfurization Efficiency
of Coal-Fired Power Plants Based on RNN**

2021년 2월

서울대학교 공학전문대학원
응용공학과 응용공학전공
서미연

석탄화력발전소
탈황공정 운전 Data를 활용한
성능예측모델 적용 연구




Prediction Model of Desulfurization Efficiency
of Coal-Fired Power Plants Based on RNN

지도교수 이종민

이 논문을 공학전문석사 학위논문으로 제출함
2021년 2월

서울대학교 공학전문대학원
응용공학과 응용공학전공
서미연

서미연의 석사 학위논문을 인준함
2021년 2월

위원장	윤용태	
위원	이종민	
위원	김국현	

초 록

석탄화력 발전소의 배기가스에는 다양한 오염물질이 포함되어 배출되며, 석탄을 연소하여 발생하는 황산화물(SO_x)의 제거를 위해 탈황공정(FGD : Fuel Gas Desulfurization)을 적용하고 있다. 탈황공정은 다양한 변수에 의해 제어되므로 황산화물 배출농도를 예측하는 것은 매우 어려우나 예측 값은 공정제어와 경보회로에 적용할 수 있으므로 유용하다.

인공신경망(ANN : Artificial Neural Network)을 기반으로 하는 딥 러닝(Deep Learning)은 산업분야에서 이미지 분류, 자연어 처리, 음성 인식 등 다양하게 활용되고 있으며 본 연구에서는 석탄화력 발전소의 운전 데이터와 인공신경망을 활용하여 배기가스 중 황산화물의 배출 농도를 예측하는 탈황공정 성능 예측 모델을 제안하였다.

시계열 데이터의 예측에는 인공신경망의 한 종류인 순환신경망(RNN : Recurrent Neural Network)이 사용되는데 본 논문에서는 NARX와 LSTM 모델을 설계하여 실험을 통해 성능을 비교 평가하고 최적화하였다. 최적화된 모델은 발전소 운전 데이터를 적용하여 검증하였고 순환신경망을 기반으로 제안된 예측 모델이 석탄화력 발전소 황산화물 배출 농도 예측에 활용될 수 있음을 보여주었다.

주요어 : 시계열 데이터 예측, FGD, NARX, LSTM, RNN

학번 : 2019-24319

목 차

I. 서론	1
1.1 연구 목적	1
1.2 연구 개요	3
1.3 논문의 구성	4
II. 배경 이론	5
2.1 공정 개요	6
2.1.1 탈황공정(Flue Gas Desulfurization)	7
2.1.2 탈황공정 화학반응식	8
2.2 Deep Learning	9
2.2.1 MLP (Multi-Layer Perceptron)	9
2.2.2 Hyper Parameter	12
2.3 RNN (Recurrent Neural Network)	17
2.3.1 NARX(Nonlinear Auto Regressive Neural Network with Exogenous Input)	18
2.3.2 LSTM(Long Short Term Memory network)	19
III. 실험 설계	21
3.1 용어의 정의	21
3.2 실험 데이터	24
3.2.1 시계열 데이터	24
3.2.2 실험 데이터 취득	27

3.2.3	상관관계 분석	29
3.3	실험방법 설계	33
3.3.1	NARX 모델 구조	34
3.3.2	LSTM 모델 구조	36
IV.	탈황예측 실험 및 결과	37
4.1	60초 데이터 탈황예측 실험	37
4.1.1	NARX 모델 최적화	38
4.1.2	LSTM 모델 최적화	40
4.1.3	60초 Event 데이터 검증	43
4.2	120초 데이터 탈황예측 실험	48
4.2.1	NARX 모델 최적화	49
4.2.2	LSTM 모델 최적화	51
4.2.3	120초 Event 데이터 검증	54
V.	탄중변화 실험 및 결과	59
5.1	60초 데이터 탄중변화 실험	61
5.1.1	NARX 모델 최적화	62
5.1.2	LSTM 모델 최적화	63
5.1.3	60초 탄중변화 데이터 예측 검증	65
5.2	120초 데이터 탄중변화 실험	68
5.2.1	NARX 모델 최적화	69
5.2.2	LSTM 모델 최적화	70
5.2.3	120초 탄중변화 데이터 예측 검증	72
VI.	결론	75

참고 문헌 78

Abstract 81

그림 목 차

그림 1.	배기가스계통 대기오염물질 저감설비 구성	6
그림 2.	탈황공정 설비 개요	7
그림 3.	Perceptron의 개요	9
그림 4.	심층신경망의 구조	11
그림 5.	RNN의 구조	17
그림 6.	NARX 모델의 구조	18
그림 7.	LSTM 모델의 구조	19
그림 8.	10~120초 샘플링 데이터의 예측 성능	24
그림 9.	10~120초 Time Sampling 데이터의 예측 성능	25
그림 10.	SOx 배출 농도에 영향을 주는 변수	27
그림 11.	상관관계 분석 결과(20초 데이터)	31
그림 12.	상관관계 분석 결과(40초 데이터)	31
그림 13.	상관관계 분석 결과(60초 데이터)	32
그림 14.	상관관계 분석 결과(120초 데이터)	32
그림 15.	최적화 모델 도출	33
그림 16.	활용된 NARX 모델의 구조	34
그림 17.	활용된 LSTM 모델의 구조	36
그림 18.	학습에 사용된 SOx 배출 농도 60초 데이터	37
그림 19.	NARX Model 60s Data Input Delay 실험 결과	38
그림 20.	NARX Model 60s Data Hidden Unit 실험 결과	39
그림 21.	NARX Model 60s Data Training Function 실험 결과	39
그림 22.	LSTM Model 60s Data Training Function 실험 결과	40

그림 23.	LSTM Model 60s Data Hidden Layer 실험 결과	41
그림 24.	LSTM Model 60s Data Learning Rate 실험 결과	41
그림 25.	LSTM Model 60s Data Epochs 실험 결과	42
그림 26.	검증에 사용된 SO _x 배출 농도 60초 Event 데이터	43
그림 27.	NARX 60s Data 최적화 모델 Event Data 예측	44
그림 28.	NARX 60s Data 최적화 모델 재학습 후 Event Data 예측	44
그림 29.	NARX 60s Data 최소화 모델 Event Data 예측	45
그림 30.	NARX 60s Data 최소화 모델 재학습 후 Event Data 예측	45
그림 31.	LSTM 60s Data 최적화 모델 Event Data 예측	46
그림 32.	LSTM 60s Data 최적화 모델 재학습 후 Event Data 예측	46
그림 33.	60s Data 최적화 모델 Event Data 검증 결과	47
그림 34.	학습에 사용된 SO _x 배출 농도 120초 데이터	48
그림 35.	NARX Model 120s Data Input Delay 실험 결과	49
그림 36.	NARX Model 120s Data Hidden Unit 실험 결과	50
그림 37.	NARX Model 120s Data Hidden Unit 실험 결과	50
그림 38.	LSTM Model 120s Data Training Function 실험 결과	51
그림 39.	LSTM Model 120s Data Hidden Layer 실험 결과	52
그림 40.	LSTM Model 120s Data Learning Rate 실험 결과	52
그림 41.	LSTM Model 120s Data Epochs 실험 결과	53
그림 42.	검증에 사용된 SO _x 배출 농도 120초 Event 데이터	54
그림 43.	NARX 120s Data 최적화 모델 Event Data 예측	55
그림 44.	NARX 120s Data 최적화 모델 Event Data 예측	55
그림 45.	NARX 120s Data 최소화 모델 Event Data 예측	56

그림 46.	NARX 120s Data 최소화 모델 재학습 후 Event Data 예측	56
그림 47.	LSTM 120s Data 최적화 모델 Event Data 예측	57
그림 48.	LSTM 120s Data 최적화 모델 재학습 후 Event Data 예측	57
그림 49.	120s 최적화 모델 Event Data 검증 결과	58
그림 50.	탄종변화 실험에 사용된 60초 데이터	59
그림 51.	탄종변화 학습에 사용된 60초 데이터	61
그림 52.	검증에 사용된 탄종변화 60초 데이터	65
그림 53.	60초 탄종변화 데이터 NARX 모델 검증결과	66
그림 54.	60초 탄종변화 데이터 LSTM 모델 검증결과	67
그림 55.	탄종변화 학습에 사용된 120초 데이터	68
그림 56.	검증에 사용된 탄종변화 120초 데이터	72
그림 57.	120초 탄종변화 데이터 NARX 모델 검증결과	73
그림 58.	120초 탄종변화 데이터 LSTM 모델 검증결과	74

표 목 차

표 1.	SO _x 배출 농도에 영향을 주는 변수	28
표 2.	SO _x 배출농도와 공정변수의 상관계수	29
표 3.	NARX 모델의 입출력 변수	35
표 4.	탄중변화 실험 데이터 소비탄 분석 결과	60
표 5.	탄중변화 실험 혼소 탄중 데이터	60
표 6.	60s Data NARX Hyper Parameter : Training Function	62
표 7.	60s Data NARX Hyper Parameter : Input Delay	62
표 8.	60s Data NARX Hyper Parameter : Hidden Unit	62
표 9.	60s Data LSTM Hyper Parameter : Training Function	63
표 10.	60s Data LSTM Hyper Parameter : Hidden Layer	63
표 11.	60s Data LSTM Hyper Parameter : Learning Rate	64
표 12.	60s Data LSTM Hyper Parameter : Epochs	64
표 13.	60초 탄중변화 데이터 최적화 모델 검증 결과	65
표 14.	120s Data NARX Hyper Parameter : Training Function	69
표 15.	120s Data NARX Hyper Parameter : Input Delay	69
표 16.	120s Data NARX Hyper Parameter : Hidden Unit	69
표 17.	120s Data LSTM Hyper Parameter : Training Function	70
표 18.	120s Data LSTM Hyper Parameter : Hidden Layer	70
표 19.	120s Data LSTM Hyper Parameter : Learning Rate	71
표 20.	120s Data LSTM Hyper Parameter : Epochs	71
표 21.	120초 탄중변화 데이터 최적화 모델 검증 결과	72

제 1 장

서 론

1.1 연구 목적

석탄화력 발전소(Coal-fired Power Plant)에서 배출되는 가스 상의 황산화물(SOx), 질소 산화물(NOx) 등은 대기에서 미세먼지 생성 반응의 주요 전구 물질로 작용한다[1]. 환경부에서는 기후변화 대응을 위해 노후 석탄화력 발전소 폐지, 신규 석탄화력 발전소 건설 제한 등 석탄연료사용 규제를 시행하고 있고 황산화물의 배출은 대기관리권역의 대기환경개선에 관한 특별법 시행에 따라 2024년까지 전망 배출량(40,239톤/년) 대비 48% 감축을 목표로 하고 있다[2][3][4].

영흥 화력 발전소는 수도권 대기관리권역에 포함되어 황산화물 배출 농도와 총량 규제를 받고 있다. 고효율 탈황공정(FGD : Fuel Gas Desulfurization)이 설치되어 있으나 석탄(Coal) 수급상황에 따라 고유황탄을 혼합하여 연소하고 최근 석회석(Limestone) 고갈로 품질이 저하되고 있어 황산화물 배출 농도가 높아질 우려가 있다[4][5][6].

황산화물 배출 농도 초과 시 과태료 부과 등 규제에 따른 손실이 발생되므로 배출 농도 예측을 통한 사전 대응이 필요하다. 탈황공정 출구 측과 연돌(Stack)에 황산화물 농도계측기가 설치되어 있으나 발생 이후 측정까지 물리적인 시간지연이 발생되므로 배출 농도를 사전에 예측할 수 있다면 다양한 방법으로 배출농도 제어에 활용할 수 있다.

국내에는 석탄화력 발전소의 황산화물 배출 농도 예측에 인공신경망 (ANN : Artificial Neural Network)의 한 종류인 순환신경망 (RNN : Recurrent Neural Network)을 적용한 사례가 없으나 해외 발전소에서는 대기 환경규제가 강화되면서 탈황공정과 탈질공정에 예측모델을 적용한 연구 사례가 있다[7][8]. 국내에서도 전력수요, 주식가격 등 다양한 분야에서 시계열 데이터를 예측하기 위해 순환신경망을 활용하는 연구가 활발히 이루어지고 있다[9][10][11]. 따라서, 본 논문은 석탄화력 발전소 탈황공정의 운전 데이터에 순환신경망을 적용한 황산화물 배출 농도 예측 모델을 제안하고자 한다.

1.2 연구 개요

시계열 데이터 예측(Time-Series Data Prediction)은 주로 회귀분석, 인공신경망(ANN)과 같은 방법을 사용해왔다[12][13]. 인공신경망의 한 종류인 순환신경망(RNN)은 시계열 데이터를 효과적으로 처리할 수 있는 구조를 가지고 있으며 순차적 데이터(Sequential Data)의 특징 추출에 적합하여 전력사용량, 날씨, 주가예측 등 다양한 분야에 사용되고 있다[9][11][14].

탈황공정의 분산제어시스템(DCS : Distributed Control System) 데이터는 일정한 표본화 간격(Sampling Interval)에 따라 변환되어 시계열 데이터로 저장된다. 시계열 데이터를 예측하려면 가까운 시간 간격(Time Step)의 공정 데이터와 멀리 떨어진 데이터 간의 영향을 반영해야 하는데 이러한 특성을 장기 의존성(Long-Term Dependencies) 이라고 한다[15]. 일반적인 순환 신경망의 경우 장기 의존성을 처리하는데 효율적이지 못하며 이는 그래디언트 소실 문제(Vanishing gradient problem)라고 알려져 있다[16]. 현재는 이러한 문제를 해결하기 위해 셀(Cell)의 상태를 기억하는 메모리 구조를 가진 LSTM(Long Short Term Memory Networks), GRU(Gated recurrent unit) 등의 모델이 제안되어 다양하게 활용되고 있다[17].

연구는 NARX(Nonlinear Auto-regressive Exogenous Neural Network) 모델과 LSTM 모델을 적용 하여 탈황공정 배기가스 황산화물 배출 농도를 예측하는 실험을 진행한다. NARX 모델은 은닉층에는 피드백이 없고 출력층에서만 피드백 구조를 가지는 구조로 딥러닝(Deep Learning) 모델인 LSTM과 예측 성능을 비교한다.

1.3 논문의 구성

본 논문의 구성은 다음과 같다.

2장에서는 석탄화력 발전소 탈황공정에 대한 배경지식을 제공하고 덤퍼닝과 순환신경망 모델에 대해 알아본다. 3장에서는 본 논문에서 사용하는 용어를 정의하고 실험 데이터와 실험 방법에 대해 설명한다.

4장에서는 앞서 설명한 데이터와 순환신경망 모델을 통해 실험을 수행하고 실제 공정의 운전 데이터로 모델의 성능을 검증한다. 5장에서는 연소되는 석탄이 변화될때 제안된 모델의 예측 성능의 검증을 위해 장기 데이터를 취득하여 실험을 한다.

마지막으로 6장에서는 연구결과를 정리하고 실제 석탄화력 발전소에 적용하기 위한 후속 연구를 제안하고자 한다.

제 2 장

배경 이론

본 장에서는 연구를 위한 배경이론 및 사전연구 사항에 대해 기술하였다. 2.1 [공정 개요]에서는 석탄화력 발전소의 배기가스를 처리하는 탈황공정에 대해 기술하였다. 2.2 [Deep Learning]에서는 본 연구에서 사용한 딥러닝 관련 이론에 대해 기술하였고 2.3 [RNN : Recurrent Neural Network]에서는 순환신경망의 대표적인 알고리즘인 NARX와 LSTM 모델에 대해 기술하였다.

2.1 공정 개요

황산화물(SOx), 질소산화물(NOx), 암모니아(NH3), 휘발성 유기 화합물(VOCs)은 초미세먼지(PM2.5)의 반응 물질로 이중 황산화물은 발전소 및 공장에서 화석 연료(주로 석탄, 석유)의 연소과정에서 발생하는 배출량이 대부분을 차지한다[6].

전국으로 확대 시행된 대기관리권역 미세먼지 총량 관리제로 인해 수도권에 위치한 석탄화력 발전소의 경우 미세먼지 발생 저감을 위해 겨울철 발전출력을 제한하는 등의 규제를 받고 있다[3]. 석탄화력 발전소는 대기환경보전법에 따라 황산화물의 배출 농도를 배출허용기준 이내로 운영해야 하기 때문에 고효율의 탈황공정을 채택하고 있으며 영흥 화력 발전소에 적용된 탈황공정은 황산화물의 제거 효율이 95% 이상이다.

보일러에서 연소된 배기가스는 그림 1과 같이 탈질공정을 통해 질소산화물을 저감하고 전기집진기를 통해 먼지를 제거한 후 탈황공정으로 유입된다. 석탄화력 발전소는 탈황공정을 통해 배출되는 황산화물을 최소화하는데 이 과정에서 공정에 적합한 온도를 유지하기 위해 열 교환기를 거치게 된다[18]. 배기가스는 각 공정을 거치면서 대기오염물질이 제거되고 열교환을 통해 온도를 높인 후 연돌(Stack)을 통해 대기로 배출된다.

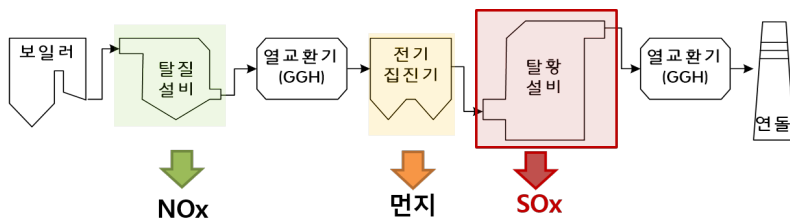


그림 1: 배기가스계통 대기오염물질 저감설비 구성

2.1.1 탈황공정(Flue Gas Desulfurization)

탈황공정(FGD : Flue Gas Desulfurization)은 석탄이 연소될 때 발생되는 배기가스에 포함된 황산화물을 석회석 슬러리(Limestone Slurry)에 흡수 반응시켜 석고 슬러리(Gypsum Slurry)로 전환한다[19].

그림 2와 같이 덕트(Duct)를 통해 탈황공정으로 유입된 배기가스는 흡수탑(Absorber) 상부로 이동하고 석회석은 분쇄되어 슬러리 상태로 흡수탑 상부에서 하부로 분사되어 배기가스 중의 황산화물과 반응하여 석고 슬러리가 생성된다[20]. 발생한 석고 슬러리는 이송되어 탈수된 후 발전 부산물(탈황 석고)로 재활용 되며 황산화물이 제거된 배기가스는 재가열되어 연돌로 배출된다[21].

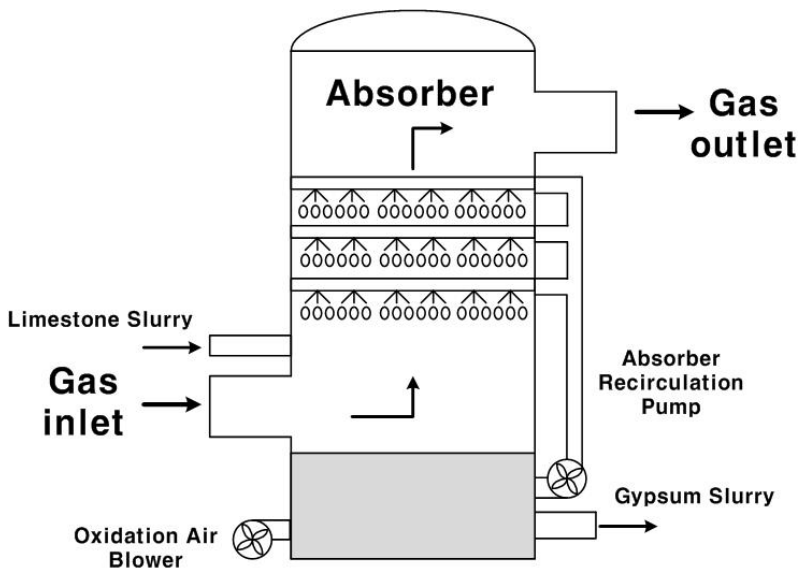


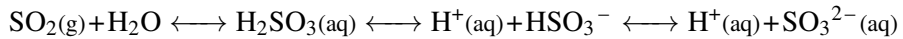
그림 2: 탈황공정 설비 개요

2.1.2 탈황공정 화학반응식

앞서 설명한 탈황공정의 화학반응식은 다음과 같다[21].

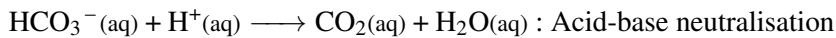
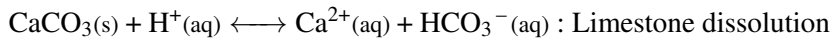
① 흡수 반응

수용성 황산화물이 흡수탑을 통과하며 분사된 물에 흡수되는 반응

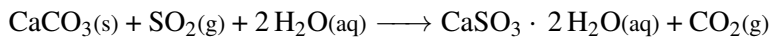


② 중화 반응

물에 흡수된 산성의 황산화물과 알칼리 흡수제인 석회석 슬러리(Limestone Slurry)가 반응하여 황산화물이 반수석고로 결합되는 반응

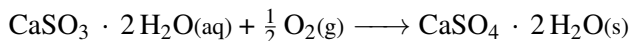


위의 흡수 반응과 중화 반응을 정리하면 다음과 같다.



③ 산화 반응

반수석고와 수용 황산화물은 전리 역반응이 쉽게 일어나므로 압축공기를 강제 공급하여 산소와 결합시키는 과정



④ 결정화 반응

황산화물과 석회석이 반응되어 생성된 석고(Gypsum)는 분포된 Seed의 성장점으로부터 길이방향으로 100~150까지 성장한다. 산화되지 않은 반수석고의 반응을 부 반응, 산화 반응까지 완료되어 역반응이 되지 않는 반응을 주 반응이라 한다.

(부 반응) $\text{CaSO}_3 \cdot \text{H}_2\text{O}$ *Sulfate*

(주 반응) $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ *Sulfate(Gypsum)*

2.2 Deep Learning

2.2.1 MLP (Multi-Layer Perceptron)

인공신경망의 기본 구조인 퍼셉트론(Perceptron)은 입력과 출력이 있는 단일 뉴런(Neuron) 구조로 되어있으며 그림3의 형태로 단순하게 표현할 수 있다[14]. 그림3에서 x 는 뉴런의 입력이고 w 는 가중치, b 는 바이어스, y 는 출력을 나타낸다.

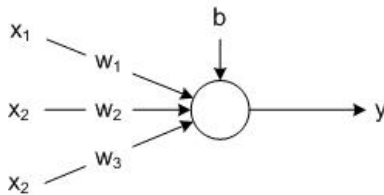


그림 3: Perceptron의 개요

n 차원 벡터가 있는 입력층을 X 벡터로 표현하고 가중치를 W 벡터로 표현하면 식 2.1을 얻을 수 있다.

$$X = (x_1, x_2, \dots, x_n)^T, W = (w_1, w_2, \dots, w_n)^T \quad (2.1)$$

식 2.2는 바이어스를 포함한 신경망 노드의 연산을 표현한 것이다. z 는 각각의 입력 값에 가중치를 곱하고 바이어스를 더하여 계산한 값으로 활성화함수(Activation Function)의 입력이 된다.

$$z = W^T X + b \quad (2.2)$$

식 2.3은 출력 뉴런의 활성화함수(Activation Function) 연산을 표현한 것으로 τ 는 활성화함수를 나타내며 y 는 출력 뉴런의 결과값이다[22].

$$y = \tau(z) \quad (2.3)$$

활성함수는 입력신호와 가중치의 연산 합을 출력 신호로 변환하는 함수이며 선형 함수도 있지만 일반적으로 비선형 함수를 사용한다.

데이터에서 복잡한 비선형 특징을 찾을 수 있어 딥러닝(Deep Learning) 훈련에 중요한 역할을 하며 본 논문에서는 LSTM 모델에서 *sigmoid* 함수와 *tanh* 함수를 사용한다.

sigmoid 함수는 식 2.4로 표현되며 출력값의 범위는 $[0,1]$ 이고 *tanh* 함수는 식 2.5로 표현되며 출력 값의 범위는 $[-1,1]$ 이다[22].

$$\tau(z) = \frac{1}{1 + e^{-z}} \quad (2.4)$$

$$\tau(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (2.5)$$

심층신경망(DNN : Deep Neural Network)을 구성하는 다층 퍼셉트론(Multi-Layer Perceptron)은 그림 4와 같이 입력층(Input Layer)과 출력층(Output Layer) 사이에 은닉층(Hidden Layer)이 있는 구조이다[23]. 은닉층을 여러 개의 층(Layer)으로 구성하면 심층신경망으로 취급되며 심층신경망을 학습시키는 알고리즘이 딥러닝(Deep Learning)이다.

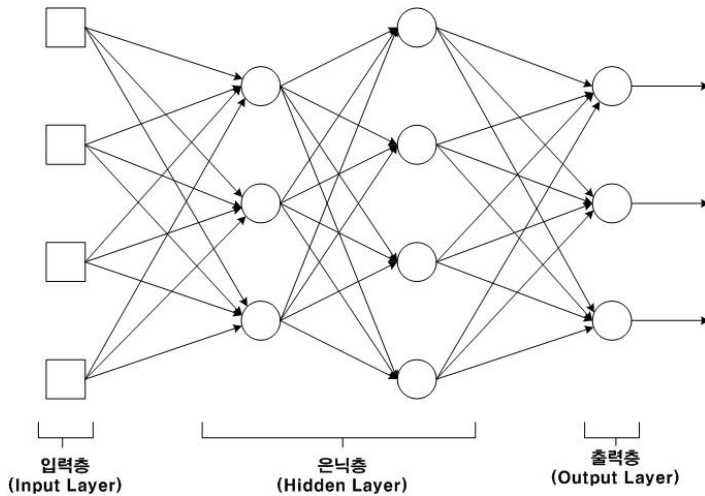


그림 4: 심층신경망의 구조

입력과 출력층은 훈련 집합이 주어지는 문제에 따라 노드(node)의 개수가 결정되는데 은닉층을 구성하는 층(Layer)의 개수, 학습 함수(Training Function)의 종류, 학습률(Learning Rate) 등은 신경망을 학습시킬 때 사용자가 지정해야 하는 값으로 이를 하이퍼 파라미터(Hyper Parameter)라 한다[24].

심층신경망을 이루는 은닉층에서는 퍼셉트론의 연산이 연속적으로 일어나며 데이터를 학습하면서 자동으로 계층적인 특징을 추출하게 되고 이렇게 추출된 특징을 활용하여 다양한 문제를 풀 수 있게 된다[22].

2.2.2 Hyper Parameter

딥러닝(Deep Learning) 모델을 학습시키기 위해 사용자가 지정해야 하는 매개변수를 하이퍼 파라미터(Hyper Parameter)라 부르며 이는 일반적인 파라미터(Parameter)와는 구분되어 사용된다. 학습을 통해 최적화해야 하는 모델 내부의 가중치가 신경망 파라미터이다[25].

모델을 학습시키기 위해 외부에서 사용자가 설정해야 하는 값은 문제에 따라 다르고 정해진 값은 없으므로 실험을 통해 하이퍼 파라미터 값을 변화하여 모델을 최적화하게 된다.

1) 학습 함수(Training Function : Optimizer, Training Algorithms)

신경망 모델의 파라미터를 조정하기 위해 학습 함수를 사용하는데 이를 최적화 함수라고 하며 목적함수 $\mathcal{J}(\Theta)$ 를 최소화하도록 신경망 파라미터를 갱신하는 오류 역전파(error backpropagation) 알고리즘이다[23].

학습 알고리즘은 확률적 경사하강 알고리즘과 변형된 최적화 함수를 사용하며 본 논문에서는 LSTM 모델의 경우 SGDM(Stochastic Gradient Descent with Momentum: 모멘텀을 사용한 확률적 경사하강법), RMSProp(Root Mean Square Propagation), Adam(Adaptive Moment Estimation)을 사용하였다[22][23].

경사하강 알고리즘은 목적함수 $\mathcal{J}(\Theta)$ 를 최소로 하는 최적해 ($\tilde{\Theta}$)를 찾기위해 그레디언트와 학습률(ρ)을 적용하는 방법으로 식2.6로 표현할 수 있으며 이 식에서 모델의 매개변수는 Θ , 목적함수는 $\mathcal{J}(\Theta)$, 학습률은 ρ 이며 $\Delta\mathcal{J}$ 는 그레디언트로 \mathbf{g} 로 나타내었다[22].

$$\Theta = \Theta - \rho\mathbf{g} \quad (2.6)$$

$$\mathbf{g} = \frac{\partial \mathcal{J}}{\partial \Theta} \quad (2.7)$$

경사하강 알고리즘은 전체 데이터 세트를 사용하여 샘플 하나의 그레이디언트를 계산하고 즉시 매개변수를 갱신한다. 이러한 연산을 모든 샘플에 수행하면 1회 epoch라고 하며 최적해 $\tilde{\Theta}$ 를 찾을때까지 epoch를 반복해서 연산한다. SGD(확률적 경사하강법) 알고리즘은 Minibatch라고 부르는 훈련 데이터의 일부를 사용하여 매개변수를 갱신하고 각 epoch 시작할 때마다 샘플의 순서를 섞어 알고리즘에 임의성을 부여하며 Minibatch 계산값이 전체 데이터 세트의 추정값이 된다[22].

SGDM(Stochastic Gradient Descent with Momentum)

확률적 경사하강법에서 최적해를 찾기 위해 그레이디언트를 추정할때 모멘텀(Momentum)항을 추가하여 수렴속도를 개선하는 방법으로 속도를 나타내는 벡터 \mathcal{V} 를 모멘텀으로 사용한다[22].

$$\mathcal{V} = \alpha \mathcal{V} - \rho \frac{\partial \mathcal{J}}{\partial \Theta} \quad (2.8)$$

$$\Theta = \Theta + \mathcal{V} \quad (2.9)$$

식 2.8에서 속도 벡터 \mathcal{V} 는 이전 그레이디언트를 누적한 것으로 볼수 있으며 α 의 범위는 [0,1]이고 이 값이 1에 가까워 질수록 이전 그레이디언트 값에 큰 가중치를 주며 $\tilde{\Theta}$ 의 값이 매끄럽게 수렴하게 된다[22].

RMSProp

학습률 값이 너무 크면 최적해를 중심으로 이동량이 너무 커서 해를 지나치는 오버슈팅 현상이 나타나게 되고 학습률이 너무 작으면 최적해까지 적은값으로 수렴하게 되어 시간이 오래 걸린다. 학습률 ρ 를 적절하게 설정하는 것은 매우 중요하므로 적응적 학습률 기법을 적용하여 매개변수

에 따라 학습률을 조정하여 사용한다. RMSProp은 오래된 그레디언트의 영향력을 지수적으로 감소시키기 위해 가중 이동 평균(Weighted moving average) 기법을 적용한다. 식 2.10에서 \mathbf{r} 은 이전 그레디언트를 누적한 벡터를 의미하며 식 2.11에서 $\Delta\Theta$ 의 값은 매개변수 Θ 를 갱신하기 위해 더해 주는 값이다[22].

$$\mathbf{r} = \alpha \mathbf{r} - (1 - \alpha) \mathbf{g} \odot \mathbf{g} \quad (2.10)$$

$$\Delta\Theta = -\frac{\rho}{\epsilon + \sqrt{\mathbf{r}}} \odot \mathbf{g} \quad (2.11)$$

Adam(Adaptive Moment Estimation)

Adam은 목적함수 Θ 의 최소값을 찾기위해 RMSProp 알고리즘에 모멘텀을 추가로 적용한 알고리즘으로 그레디언트 누적 벡터 \mathbf{r} 을 계산하고 모멘텀을 적용하기 위해 속도 벡터 \mathcal{V} 를 계산한뒤 각각의 α 값에따라 Θ 를 구하는 방법이다. 식 2.12는 $\Delta\Theta$ 의 계산식을 표현하였으며식 2.10에서 \mathbf{r} 과 식 2.8에서 \mathcal{V} 의 계산 식을 적용할 수 있다[22].

$$\Delta\Theta = -\frac{\rho}{\epsilon + \sqrt{\mathbf{r}}} \mathcal{V} \quad (2.12)$$

NARX 모델의 경우 본 논문에서 `trainlm`, `trainrp`, `traingdx` 함수를 사용하였는데 이 함수들은 다중퍼셉트론을 학습하기 위한 최적화 알고리즘으로 다양한 오류 역전파(error backpropagation) 알고리즘을 사용하고 있다. 각 함수와 사용하는 알고리즘을 소개하면 `trainlm`은 Levenberg-Marquardt backpropagation, `trainrp`는 Resilient backpropagation, `traingdx`는 Gradient descent with momentum and adaptive learning rate backpropagation이다.

2) 은닉층(Hidden Layer)의 개수

심층신경망(DNN)은 은닉층(Hidden Layer)의 개수를 늘려 복잡한 문제를 풀 수 있지만 층(Layer)을 무조건 늘린다고 모델의 성능이 좋아지는 것은 아니다[15]. 은닉층이 많아 신경망이 제대로 학습되지 않으면 오히려 성능이 나빠질 수 있으므로 실험을 통해 적절한 값을 찾아야 한다[23]. NARX 모델의 경우는 은닉층이 단층으로 구성되므로 은닉 유닛(Hidden Unit) 수를 변화시켜 최적화 한다.

3) Adaptive Learning Rate(적응적 학습률)

학습률(Learning Rate)은 앞서 학습 함수에 의해 신경망 파라미터를 갱신하는 정도를 결정하며 실험과정에서 학습률 값을 변화시키며 모델 성능을 평가하여 결정한다. 학습률 값은 훈련(Training)이 끝날 때까지 고정된 값을 사용할 수도 있지만 훈련을 시작할 때는 상대적으로 큰 학습률을 선택하고 일정한 Epochs 마다 학습률 값을 점진적으로 줄이는 등 학습 상황에 따라 조절하여 사용할 수 있다. 이렇게 하면 훈련 초기에는 변화를 크게하여 전체 학습 시간이 단축되고, 훈련이 진행되어 오차(Error)가 최소 값에 수렴할수록 신경망 파라미터를 더 적게 변화시킬 수 있다[22].

다양한 방식으로 모델의 학습에 적용할 수 있지만 본 논문에서는 학습률을 설정하고 최대 Epochs를 적절한 크기로 나누어 앞부분에서는 학습률을 그대로 적용하고 뒷부분에는 학습률에 0.2를 곱하여 적용되는 값을 지속적으로 감소시키는 방법으로 실험 하였다.

4) Epochs

최적화 함수를 통해 은닉층의 가중치인 신경망 파라미터를 갱신하게 되는데 전체 학습데이터(Training Data Set)를 완전히 한 번 통과하는 것을 1 Epoch라고 하며 최대 Epochs 횟수를 지정하여 반복 횟수를 설정할 수 있다. 검증 데이터(Validation Data Set)를 별도로 정하는 경우 훈련을 통한 파라미터의 성능을 평가하여 정해진 Epochs을 모두 수행하지 않아도 학습을 종료하도록 설정할 수 있으며 신경망 파라미터가 과적합되어 훈련 성능이 저하되는 것을 막을 수 있다. 본 논문에서는 최적화 모델을 도출하기 위해 Epochs 수를 변화하여 예측 성능을 비교하는 실험을 하였고 검증 데이터를 별도로 정하지 않고 정해진 Epochs 수를 모두 수행하도록 하였다.

2.3 RNN (Recurrent Neural Network)

순환신경망은 인공지능의 한 종류로 과거의 정보가 미래에 영향을 줄 수 있는 구조로 확장된 신경망이다[15]. 정보를 지속적으로 전달하는 구조로 이전 시간 스텝($t - 1$)의 상태(State, $h(t - 1)$)를 입력으로 받아 현재(t)에 영향을 줄 수 있어 시계열 데이터 처리에 좋은 성능을 보인다.

순환신경망 셀은 그림5 와 같이 입력(x)을 받아 출력(y)을 내보내고 순환 구조를 통해 셀의 정보(h)가 신경망의 한 단계에서 다음 단계로 전달되도록 한다. 순환 신경망은 동일한 신경망이 연결된 형태로 생각할 수 있는데, 정보가 전달되는 구조가 순차적 데이터 구조의 특성을 잘 반영한다는 것을 알 수 있다.

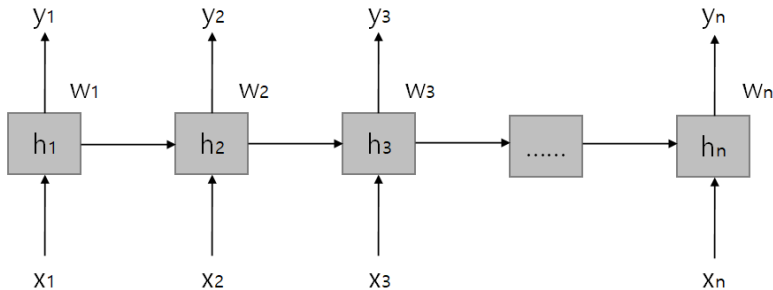


그림 5: RNN의 구조

순환신경망을 시계열 데이터에 적용하여 미래를 예측하는 연구는 많이 있으나 기존의 순환신경망은 은닉층이 늘어남에 따라 오차 역전파(Back-propagation)시 기울기 손실이 있다[16]. 이 문제를 해결하기 위해 변형된 순환신경망 모델인 LSTM(Long Short-Term Memory) 모델이 제안되었다[17]. 본 논문에서는 실험을 위해 단층 순환신경망 모델인 NARX와 다층 순환신경망 모델인 LSTM을 적용하였다.

2.3.1 NARX(Nonlinear Auto Regressive Neural Network with Exogenous Input)

NARX 모델은 순환신경망의 한 종류이며 다른 순환 신경망과는 달리 은닉층이 아닌 출력층에서만 피드백 구조를 가지는 신경망이다[9].

NARX 모델의 연산결과는 식 2.13과 같이 표현할수 있는데 출력 $y(t)$ 는 예측하고자 하는 값이고 입력은 $y(t-1)$ 과 외부 입력(Exogenous Input) $u(t-1)$ 으로 각 입력의 과거 값을 받는다. 모델의 구조에 따라 $u(t)$ 값도 참조할수 있으나 본 논문에서는 $(t-1)$ 이전의 값만을 참조하여 예측하는 구조로 설계하였으며 NARX 모델의 입력이 몇 Time Step 이전 값까지 참조하는지를 결정해야 하는데 이 값을 Input Delay라 한다.

$$y(t) = F[y_{t-1}, y_{t-2}, y_{t-3}, \dots, u_t, u_{t-1}, u_{t-2}, u_{t-3}, \dots] + \epsilon_t \quad (2.13)$$

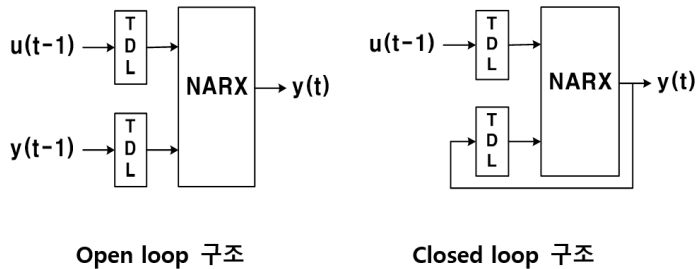


그림 6: NARX 모델의 구조

NARX 모델의 입출력 구조는 그림 6과 같이 나타낼 수 있으며 학습 시에는 Open loop 구조로 파라미터를 학습시킨 후 Closed loop 구조를 통해 Multi step 예측을 수행한다. 실제공정에서는 예측 후 제측기에 의해 실측 값을 얻을 수 있어 Multi step 예측이 불필요 하고 정확도도 낮으므로 본 논문에서는 Single step 예측을 하기 위해 학습과 테스트 모두 open loop 구조를 사용하였다.

2.3.2 LSTM(Long Short Term Memory network)

순환신경망의 장점 중 하나는 과거의 정보를 지속적으로 전달할 수 있게 해준다는 점이다. 그러나 기본적인 순환신경망은 정보가 길어지는 과정에서 이전 정보와의 관계를 학습할 수 없으므로 긴 시퀀스 데이터를 다루기 적합하지 않다[26]. 이것을 기울기 소실 (Banishing gradient problem)에 의한 데이터의 장기 의존성 문제(The problem of long-term dependency)라고 하는데 LSTM(Long Short Term Memory network)은 기본 순환신경망과는 달리 고유의 메모리 셀(Cell) 구조를 가지고 있으며 셀의 상태 값을 추가하여 정보의 기억과 전달을 통해 장기 의존성 문제를 해결하였다[16]. LSTM Cell은 망각게이트(Forget gate), 입력게이트(Input gate) 및 출력게이트(Output gate) 구조로 이루어져 있으며 각 게이트에 의해 연산된 결과를 출력한다[17]. LSTM 모델은 이전 단계의 정보를 기억하는 메모리를 가지고 있으며 그림 7에 구조를 나타내었다[27].

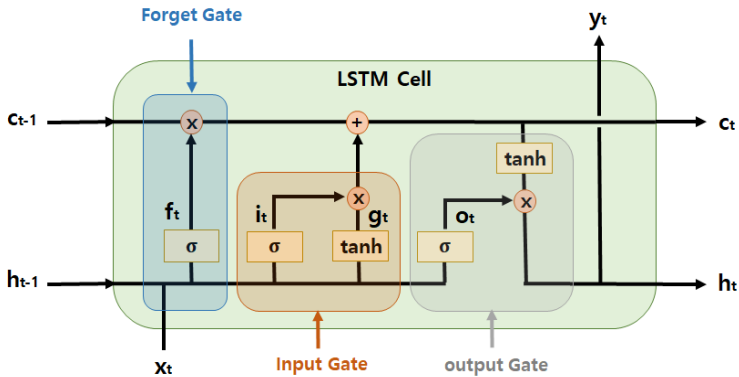


그림 7: LSTM 모델의 구조

그림 7은 식 2.14 ~ 2.19로 표현할수 있으며 \mathbf{W} 는 가중치, \mathbf{b} 는 편향 (Bias) 값으로 아래 첨자는 각각의 게이트를 표현하며 기능은 다음과 같다.

$$f_t = \text{sigmoid}(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (2.14)$$

$$i_t = \text{sigmoid}(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (2.15)$$

$$g_t = \text{tanh}(W_{gx}x_t + W_{gh}h_{t-1} + b_g) \quad (2.16)$$

$$o_t = \text{sigmoid}(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (2.17)$$

$$c_t = c_{t-1} \otimes f_t + g_t \otimes i_t \quad (2.18)$$

$$h_t = \text{tanh}(c_t) \otimes o_t \quad (2.19)$$

1) 망각게이트(forget gate : f_t) 이전 단계의 상태 값 h_{t-1} 과 새로운 입력 값 x_t 을 연산하고 활성화 함수(Activation Function)의 한 종류인 시그모이드(sigmoid)를 이용하여 셀(Cell)에 저장할 정보와 버릴 정보를 $[0,1]$ 의 범위에서 결정한다. (0: 정보를 버림, 1: 정보를 반영)

2) 입력게이트(input gate : i_t, g_t) 셀에 입력될 정보를 결정하는 단계로 sigmoid 함수를 이용하여 새롭게 입력할 정보 i_t 를 결정하고 tanh 함수를 이용하여 셀의 값을 업데이트 하기 위해 사용하는 입력 예정 값 g_t 를 결정한다.

3) 메모리셀(Memory Cell : c_t) 과거 셀의 값 c_{t-1} 과 셀의 입력 예정 값 g_t 의 연산을 통해 현재 셀의 값 c_t 를 업데이트한다.

4) 출력게이트(output gate : o_t, h_t) 출력을 결정하는 단계로 sigmoid 함수를 이용하여 출력할 값 o_t 를 계산하고, tanh 함수를 통해 $[-1,1]$ 범위의 값으로 변환된 셀의 값 c_t 와 곱하여 셀의 새로운 출력 h_t 를 결정한다.

제 3 장

실험 설계

본 장에서는 모델을 최적화 하고 성능을 비교하기 위해 필요한 실험 데이터 및 실험 방법에 대해 기술하였다.

3.1 용어의 정의

본 연구에서 순환신경망 모델을 학습시키기 위한 실험환경을 구현하고 성능을 평가하기 위해 사용한 용어를 다음과 같이 정의한다.

1) 샘플링 데이터(Sampling Data)

탈황공정은 설비의 운전을 위해 다양한 계측기가 설치되어 있으며 계측기가 취득한 아날로그(Analog) 측정값은 분산제어시스템(distributed control system, DCS)에 1초 시간간격으로 저장된다. 본 논문에서는 DCS에 저장된 데이터를 공정 모니터링 시스템을 이용하여 실험에 필요한 표본화 시간간격으로 취득하였는데 이를 샘플링 데이터라 명칭하였다.

2) 학습 데이터(Training Data & Test Data)

모델을 학습시키기 위해서는 학습 데이터가 필요하며 본 논문에서는 모델을 최적화 시키기 위해서 60초 샘플링 데이터와 120초 샘플링 데이터로 나누어 동일 기간의 연속된 데이터를 취득하였다. 이 데이터의 90%는 Training Data로 모델의 학습에 사용하고 10%는 Test Data로 학습된 모델의 성능검증에 사용하였다. 모델을 최적화 시키기 위해 사용한 데이터를 학습 데이터라 명칭하였고 본 논문에서는 Training Data set : Test Data set = 9 : 1로 설정하여 실험하였다.

3) 최적화 & 최적화 모델 & Event Data

학습 데이터를 NARX와 LSTM 모델에 대해 적용하여 높은 예측 성능 값을 나타내는 하이퍼 파라미터를 찾는 과정을 최적화라 명칭하였다. 성능평가 기준에 따라 최적화된 하이퍼 파라미터를 적용하여 훈련시킨 모델을 최적화 모델이라하고 성능 검증 시 최적화 모델을 사용하였다.

60초와 120초 데이터 각각에 대해 최적화 모델을 도출하였고 실험을 통해 최적화 된 모델이 공정의 황산화물 배출 농도를 예측할수 있는지 검증하기 위해 급격하게 변동하는 특정 기간의 8시간 데이터를 취득하였으며 이를 Event Data라고 명칭하였다.

4) 재학습 & 최소화 모델

동일한 석탄화력 발전소의 운전데이터를 사용하였지만 Event 데이터(2017.08.26, 12:00~20:00)를 취득한 시점과 최적화를 위한 학습데이터(2020.08.27, 00:00~12:00)를 취득한 시점 간에 3년의 시간 차가 있다. 학습된 최적화 모델을 Event 데이터에 적용하여 검증 결과를 도출하였으며 동기간의 학습데이터와 예측 성능을 비교하기 위해

(1)최적화 모델의 하이퍼 파라미터 설정은 유지하고 Event 데이터보다 앞선 시점의 12시간 연속 데이터(2017.08.26, 00:00~12:00)를 취득하여 다시 학습시킨 후 Event 데이터를 예측하였다. (1)의 실험방식을 재학습이라 표현하였으며 하이퍼 파라미터는 동일하나 재학습을 하면 모델 내부의 신경망 파라미터 값은 훈련에 의해 변경된다.

(2)NARX 모델을 하이퍼 파라미터의 값을 매우 낮게 재설정 하고 Event 데이터의 바로 이전 시점의 12시간 연속 데이터를 재학습 시켜 예측 성능을 평가 하였다. (2)의 실험방식으로 학습된 모델을 최소화 모델이라 명칭하였으며 LSTM 모델의 경우 예측 성능이 나빠지므로 본 논문에서는 NARX 모델의 실험결과에만 적용하였다.

5) 탄중변화 실험 데이터

석탄화력 발전소는 석탄을 혼합하여 사용하며 매일 혼합되는 석탄의 종류와 양이 달라지게 되고 8시간 이상 사용가능한 석탄을 저장할 수 있는 사일로를 통하여 석탄이 공급되므로 석탄의 연소시점을 정확하게 아는 것은 어렵다. 따라서 탄중변화에 따른 예측 성능의 정확도를 평가하기 위해서는 1일 이상의 장기간의 예측 성능평가가 필요하므로 6일(144시간)의 운전데이터를 취득하여 탄중 실험 데이터(2021.01.11 00:00~01.17 00:00)로 사용하였다.

탄중변화 실험 데이터는 탈황예측 실험 데이터와 동일하게 60초와 120초 샘플링 데이터로 나누어 취득하였다. 앞부분 3일의 운전 데이터(2021.01.11 00:00~01.14 00:00)를 통해 LSTM과 NARX 모델 각각 이전 실험과 동일한 방식으로 최적화 모델을 도출한 다음 뒷부분 3일의 운전 데이터(2021.01.14 00:00~01.17 00:00)를 최적화 모델에 적용하여 예측 성능을 검증하는 과정으로 실험을 진행하였다.

3.2 실험 데이터

3.2.1 시계열 데이터

시계열(Time series) 데이터는 공정에서 측정 장치에 의해 저장되는 연속적인 값이다. 본 연구의 목표는 공정 운전 데이터를 이용하여 주어진 시간 t 에서 샘플링 시간 T 이후의 값, 즉, 시간 $t + T$ 에서의 값을 예측하는 것이다. 샘플링 시간을 10~120초로 변화하여 취득한 실험 데이터를 LSTM 모델로 예측한 결과를 그림 8과 같이 그래프로 표현하였다.

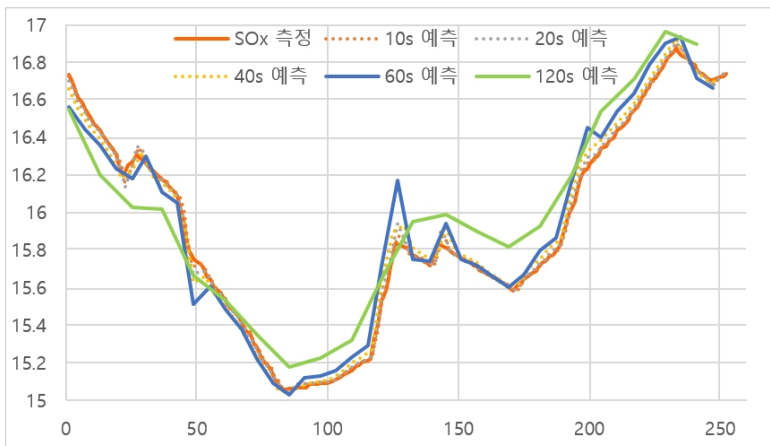


그림 8: 10~120초 샘플링 데이터의 예측 성능

10~40초 샘플링 데이터에서 예측 성능이 좋고 60초보다 샘플링 간격이 큰 데이터에서는 측정값과 편차가 발생한다. 공정의 시정수는 (Time Constant)는 60초 이므로 공정에 적용하기 위해서는 60초 이상 시간간격 이후의 값을 예측할 수 있어야 유의미하게 활용할 수 있다. 연돌(Stack)에 설치된 대기 TMS 분석기에서 측정되는 황산화물 배출 농도값은 120초의 시간 지연이 있기때문에 예측에 사용할 데이터의 샘플링 시간은 60초와 120초로 선정하여 데이터를 취득하였다.

실험에 사용한 황산화물 배출 농도값은 흡수탑 후단에 설치된 SOx 농도 계측기를 통해 실시간으로 측정하여 공정예측용 데이터로 활용되며 60초 예측값은 석회석 슬러리의 유량제어를 위한 선형 제어값과 정보회로로 구성하여 공정에 적용할 수 있다. 120초 예측값은 배출농도감시에 사용되는 대기 TMS 분석기 측정값에 적용할 수 있어 유용한 값이다.

10~120초 샘플링 데이터의 예측 성능을 그림 9에서 확인할 수 있다. 그림 8에서 예측 값과 측정값의 편차를 확인한 것과 같이 10 ~ 40초 샘플링 데이터의 RMSE 값이 60초와 120초의 RMSE 값보다 작은 것을 알 수 있으며 MAPE 값도 동일한 추세를 보인다.

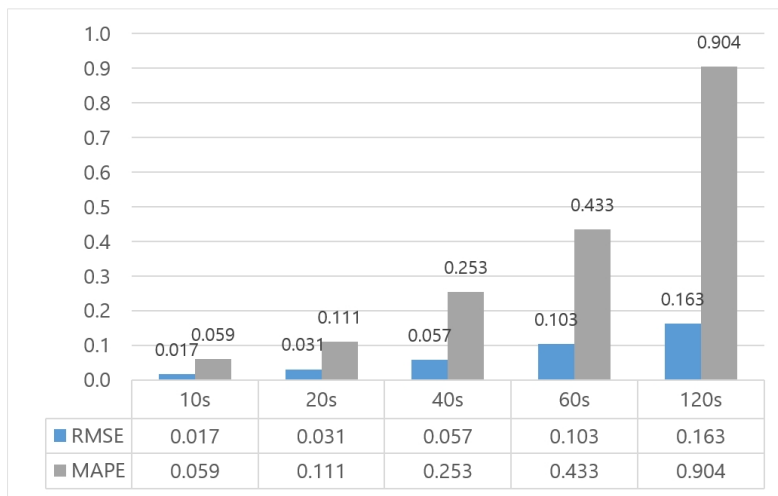


그림 9: 10~120초 Time Sampling 데이터의 예측 성능

본 논문에서 적용한 NARX와 LSTM 모델은 시계열 데이터를 학습하여 예측을 수행할 수 있으며 공정에서 측정된 입력을 통한 Single time step 예측 외에 예측된 값을 피드백 받아 Multi time step 예측도 가능하다.

탈황공정은 황산화물 배출 농도 측정값이 존재하므로 정답이 주어지는 지도 학습(Supervised Learning)에 해당하며 예측 결과를 측정값과 바로 비교할 수 있다[10]. Multi time step 예측은 본 논문의 연구 목표가 아니며 측정 데이터로 모델의 파라미터를 Update 하는 것이 Multi time step 예측보다 성능이 좋으므로 Single time step 예측을 적용하여 실험하였다.

LSTM 모델은 학습 데이터(Training Data) 전체가 예측 값에 영향을 주며 영향의 정도는 학습에 의해 결정되며 신경망 파라미터로 저장된다. NARX 모델의 경우는 영향을 줄 이전 Time step의 개수를 하이퍼 파라미터로 설정해주어야 하며 설정에 따라 모델의 성능이 달라진다. 이를 지연시간(Input Delay)라하며 입력의 시간을 몇 스텝 전까지 참조 했는지를 나타내는 것으로 NARX 모델에서 중요한 하이퍼 파라미터 이다.

3.2.2 실험 데이터 취득

학습데이터는 예측 데이터의 활용측면과 공정특성에 따라 60초와 120초의 시간간격 (Sampling Time)으로 취득하기로 하였으며 학습데이터를 취득하기 위해 탈황공정의 황산화물(SOx) 배출 농도에 영향을 미치는 변수들을 그림 10에 나타내었다[19]. 탈황공정은 실제로 더 많은 변수들에 의해 영향을 받고 있지만 선행연구와 설치된 계측기를 참고하여 데이터를 취득할 수 있는 변수를 선정하였다[19][21].

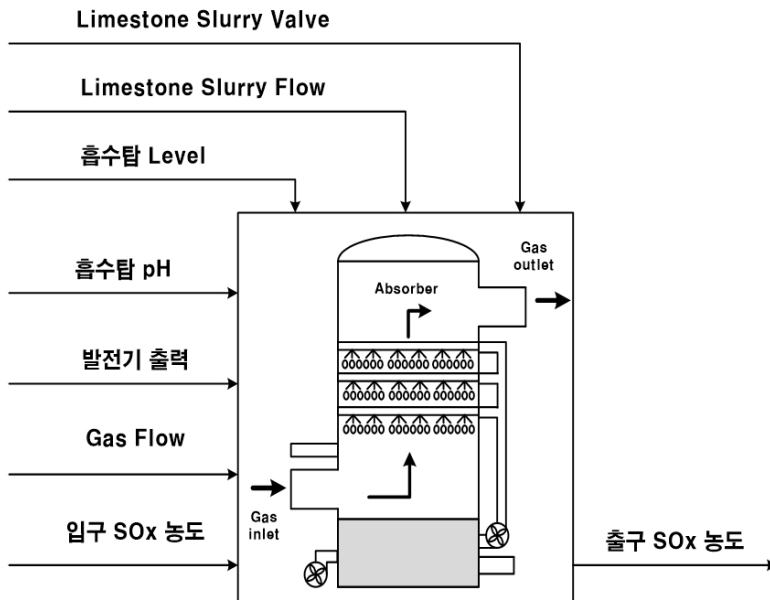


그림 10: SOx 배출 농도에 영향을 주는 변수

석탄화력 발전소 운전 데이터는 분산제어시스템(DCS)에 저장되어 공정 제어에 사용되며 SOx 배출 농도에 영향을 주는 변수들의 데이터는 공정 감시(Monitoring) 설비를 통해 취득하였다.

예측 모델 최적화에 사용된 학습 데이터의 범위는 표 1과 같으며 과적합(Over-fitting)을 방지하고 일반화(Generalization)를 만족하기 위해 충분한 기간을 선정하여 취득하였다. 학습 데이터에 대한 설명은 3.1 [용어의 정의]에 좀더 자세하게 설명되어 있다.

표 1: SOx 배출 농도에 영향을 주는 변수

변수	기호	단위	범위
출구 SOx 농도	SOx Out	ppm	15.6~17.3
입구 SOx 농도	SOx In	ppm	215.9~279.9
발전기 출력	MW	MW	708~806
흡수탑 pH	pH	pH	4.6~5.2
흡수탑 Level	Level	m	6562~6684
배기 Gas Flow	Gas Fw	m ³ /h	2344~2707
Limestone Slurry Flow	LS Fw	m ³ /h	8.2~18.9
Limestone Slurry valve	LS Vv	%	27.1~54.8

3.2.3 상관관계 분석

예측 모델에 적용할 공정 변수를 결정하기 위해 예측하고자 하는 목표 값(SOx 배출 농도)과 각 공정 변수간의 상관관계 분석을 하였다. 두 변수간의 관련성을 구하기 위해 보편적으로 이용하는 것은 피어슨 상관 계수(Pearson correlation coefficient)로 표현 식은 다음과 같다[28].

$$r = \frac{X와 Y가 함께 변하는 정도}{X와 Y가 각각 변하는 정도} \quad (3.1)$$

$$r_{xy} = \frac{\sum_i^n (X_i - \bar{X}) * (Y_i - \bar{Y})}{\sqrt{\sum_i^n (X_i - \bar{X})^2} * \sqrt{\sum_i^n (Y_i - \bar{Y})^2}} \quad (3.2)$$

상관계수 표현식 3.2에서 X_i 와 Y_i 는 데이터의 개별 샘플을 의미하며 데이터의 샘플 크기는 n 이고 i 로 인덱싱 되어 표현된다. 식에서 \bar{X} 와 \bar{Y} 는 각 샘플의 평균을 의미하며 데이터는 쌍으로 이루어진다.

표 2: SOx 배출농도와 공정변수의 상관계수

공정변수	20s data	40s data	60s data	120s data
MW	-0.330	-0.187	-0.221	0.025
SOx In	0.193	-0.292	-0.213	0.036
pH	-0.938	-0.726	-0.901	-0.498
Level	-0.908	-0.241	0.140	-0.129
Gas Fw	-0.468	-0.147	0.290	0.000
LS Fw	0.277	0.414	0.306	0.504
LS Vv	0.832	0.642	0.745	0.706
SOx Out(t-1)	0.998	0.985	0.982	0.886

표 2의 상관관계 분석 결과를 보면 SOx 배출 농도의 측정값(SOx Out(t))은 자기 자신의 이전 값(SOx Out(t - 1))과 가장 상관관계가 강하므로 자기상관성 (Auto-correlation)이 높은 데이터이다. 자기상관성(Auto-correlation)은 시계열 데이터의 현재 값(t)과 자신의 과거 값(t - 1) 사이의 상관관계를 의미한다.

20초 샘플링 데이터의 상관관계 분석결과는 pH와 흡수탑 Level 값과도 강한 음의 상관관계를 갖으며 60초 샘플링 데이터까지도 유사한 결과를 보여준다. 그러나 120초 샘플링 데이터의 분석 결과에서 대부분의 공정 변수는 상관계수가 0.5 이하로 데이터의 샘플링 시간이 길어질수록 SOx 배출 농도와 상관관계가 낮아지는것을 확인할수 있다.

120초 샘플링 데이터의 상관관계 분석에서 가장 강한 양의 상관관계를 보여준 SOx 배출 농도 이전값은 (SOx Out(t - 1)) NARX와 LSTM 모델에서 예측을 위한 입력 값으로 적절하다. 두번째로 높은 상관계수 값을 나타낸 Limestone Slurry valve 개도 값(LS Vv)은 NARX 모델의 외생 입력으로 적용할 수 있으며 60초와 120초 이후의 값을 예측하기 위한 공정변수로서 적합하다. 40초이하의 짧은 샘플링 데이터의 값을 예측할 때에는 상관관계 분석 결과에 따라 높은 상관계수 값을 나타낸 다른 공정변수를 추가로 적용해 볼수 있을 것이다.

그림 11 ~ 그림 14에서는 앞서 표 2에 정리된 상관관계 분석 결과를 산점도 그래프를 통해 확인할 수 있다. 데이터는 동일 시점의 공정 데이터로부터 각 샘플링 시간에 따라 취득된 데이터를 사용하였으며 분석에 사용된 샘플 수는 100개이다.

SOx O(t), MW_1, SOx In, pH, Level, Gas Fw, LS Fw, LS Vv, SOx O(t-1)의 산점도 행렬
 Pearson 상관 계수에 대한 95% CI

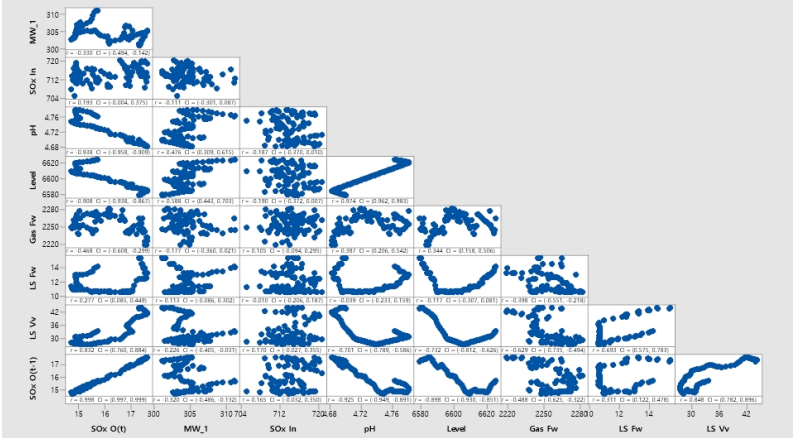


그림 11: 상관관계 분석 결과(20초 데이터)

SOx O(t), MW, SOx In, pH, Level, Gas Fw, LS Fw, LS Vv, SOx O(t-1)의 산점도 행렬
 Pearson 상관 계수에 대한 95% CI

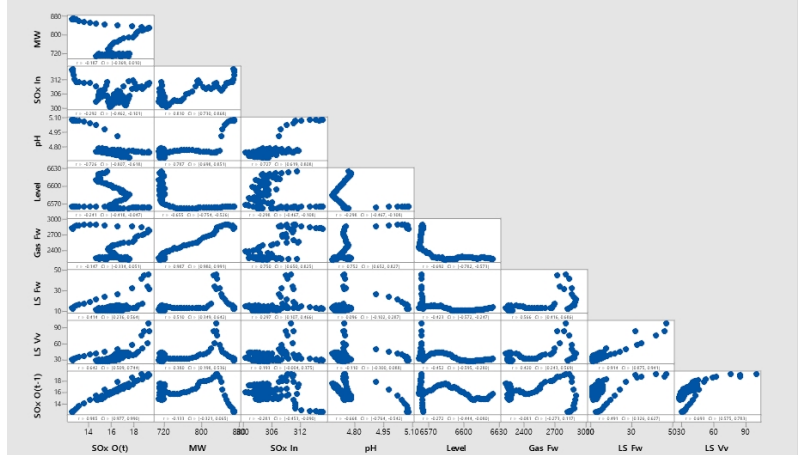


그림 12: 상관관계 분석 결과(40초 데이터)

SOx O(t), MW, SOx In, pH, Level, Gas Fw, LS Fw, LS Vv, SOx O(t-1)의 산점도 행렬
Pearson 상관 계수에 대한 95% CI

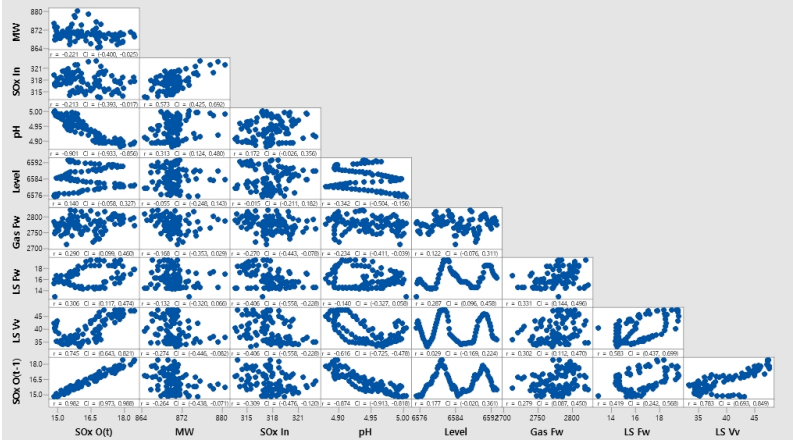


그림 13: 상관관계 분석 결과(60초 데이터)

SOx O(t), MW, SOx In, pH, Level, Gas Fw, LS Fw, LS Vv, SOx O(t-1)의 산점도 행렬
Pearson 상관 계수에 대한 95% CI

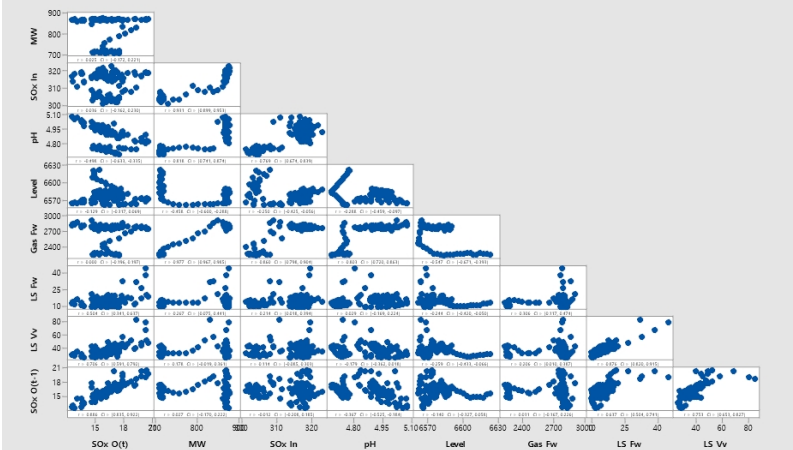


그림 14: 상관관계 분석 결과(120초 데이터)

3.3 실험방법 설계

지금까지 실험을 위한 학습 데이터를 취득하는 과정을 설명하였다. 최적화 모델을 도출하기 위한 과정은 모델 선택, 데이터 수집, 학습을 통한 모델 최적화, 예측 데이터 검증의 네 단계로 구성된다. 이 과정을 그림 15와 같이 표현할 수 있다.

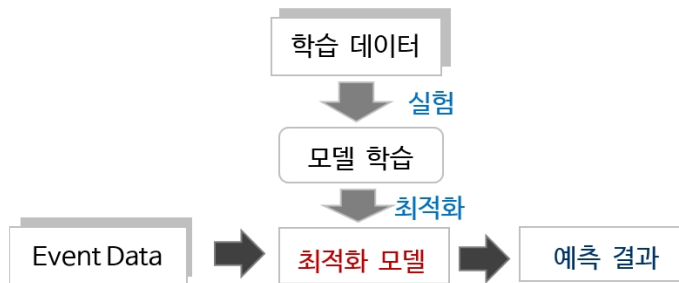


그림 15: 최적화 모델 도출

본 논문에서는 NARX와 LSTM 두개의 모델에 대해 동일한 데이터를 Training set(90%)과 Test set(10%)으로 나누어 학습시킨 뒤 성능을 평가하였다. 모델의 성능 지표는 Test set의 실제 값과 예측 값의 RMSE(Root Mean Squared Error)값과 MAPE(Mean Absolute Percentage Error) 값을 계산하여 사용하였다[19].

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\tilde{y}_i - y_i)^2}{n}} \quad (3.3)$$

$$MAPE = \frac{1}{n} * \sum_i^n \left| \frac{\tilde{y}_i - y_i}{y_i} \right| * 100\% \quad (3.4)$$

일반적으로는 RMSE 값으로 모델의 성능을 평가하지만 데이터의 절대 값이 큰 경우 Error도 크게 표현될 수 있어 최저 값이 상이한 경우 MAPE 값을 성능평가 지표로 우선하였다.

3.3.1 NARX 모델 구조

NARX(Nonlinear Auto Regressive Neural Network with Exogenous Input)모델은 순환 신경망의 한 종류로 은닉층은 단층으로 구성하였다. 하이퍼 파라미터로 뉴런 개수(Hidden Unit)와 입력 값의 지연 시간(Input Delay)을 결정할 수 있다[9][11]. 뉴런 수와 지연 시간은 해결하고자 하는 문제에 따라 다르므로 이상적인 개수가 정해져 있지 않다.

처음에는 뉴런 개수를 NARX 모델의 기본 설정 값 또는 최소 값으로 설정하고 지연 시간 또한 모델의 기본 값 으로 설정하여 예측 성능에 따라 하이퍼 파라미터를 변화시켜 최적화 한다. 실험을 위한 NARX 모델의 구조는 그림 16과 같다[7].

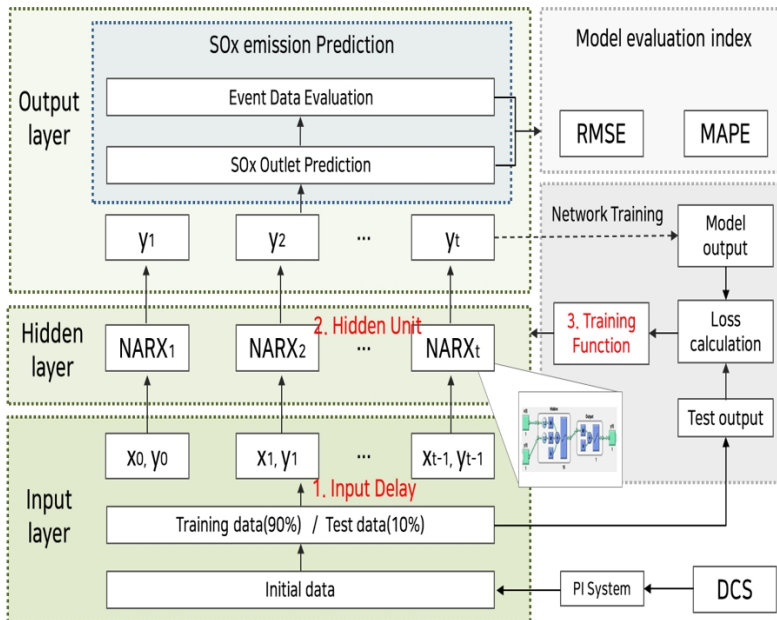


그림 16: 활용된 NARX 모델의 구조

그림 16에 표현된 NARX 모델의 입출력은 표 3과 같다. 모델의 입력은 앞서 상관관계 분석 결과에 따라 Limestone Slurry valve 개도 값과 SOx 농도 측정값이다.

표 3: NARX 모델의 입출력 변수

Exogenous Input($t - 1$)	Output(t)
Limestone Slurry valve (%)	SOx emission (<i>ppm</i>)
Input($t - 1$)	
SOx emission (<i>ppm</i>)	

NARX 모델에서는 3개의 하이퍼 파라미터를 사용하였는데 먼저 지연시간(Input Delay)은 입력을 몇 시간 스텝 전까지 참조 했는지를 나타내고 2~50으로 변화시켰다. 두번째 은닉 층의 뉴런 개수 (Hidden Unit)는 은닉층이 단층으로 되어있으므로 가중치를 갖는 뉴런의 개수이며 2~40개로 변화시켜 실험하였다. 세번째 학습 함수(Training Function)은 모델 내부의 가중치인 파라미터를 변경하는 학습 알고리즘을 말하는데 3개의 학습 함수(Trainlm, trainrp, traingdx)를 선정하여 성능을 비교하였다.

실험은 학습 데이터를 NARX 모델에 적용하여 학습시키고 이때 세개의 하이퍼 파라미터를 변화하며 예측 성능을 비교하여 최적화 모델을 도출한 다음 Event Data로 예측 성능을 검증하였다.

3.3.2 LSTM 모델 구조

표 2 상관관계분석 결과에 따르면 SOx 배출 농도 값은 높은 자기상관성을 갖는다는 것을 알 수 있으므로 LSTM 모델의 입력은 SOx 배출 농도의 이전 시간 값으로 하였다.

실험에서 사용하는 모델의 구조를 그림 17에 나타내었는데 하이퍼 파라미터는 4개로 학습 함수(Training Function : Adam, RMSProp, SGDM), 은닉층의 Layer 개수(1~300), 학습률(0.005~0.03), Epochs(80~400)이다[7]. 이에 대한 설명은 2.2.2 [Hyper Parameter]에서 자세히 기술하였다. 3.3.1 [NARX 모델 구조]의 설명과 동일하게 실험을 수행하였고 4장과 5장에서 실험 결과를 기술하였다.

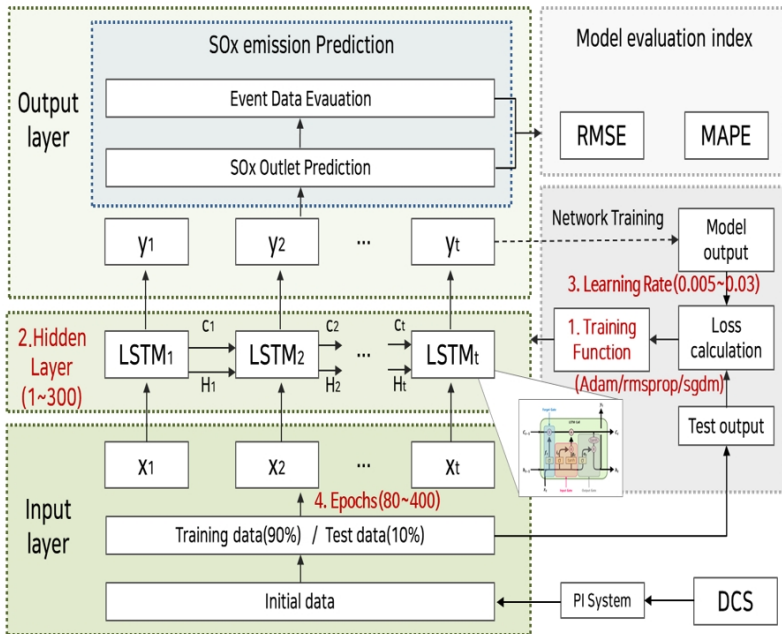


그림 17: 활용된 LSTM 모델의 구조

제 4 장

탈황예측 실험 및 결과

본 장에서는 NARX와 LSTM 모델의 하이퍼 파라미터별 예측 성능을 비교하여 최적화된 모델을 도출한 다음 석탄화력 발전소의 Event Data를 통해 예측모델의 성능을 검증하였다.

4.1 60초 데이터 탈황예측 실험

모델을 최적화하기 위해 60초 샘플링 간격으로 황산화물(SO_x) 배출 농도 값과 Limestone Slurry valve 개도 값을 취득하였다. 학습에 사용된 SO_x 배출 농도 데이터는 그림 18과 같다. 학습 데이터(2020.08.27, 00:00~12:00)는 과적합(Over-fitting)을 방지하고 일반화(Generalization)를 만족하기 위해 충분한 데이터를 취득하였다.

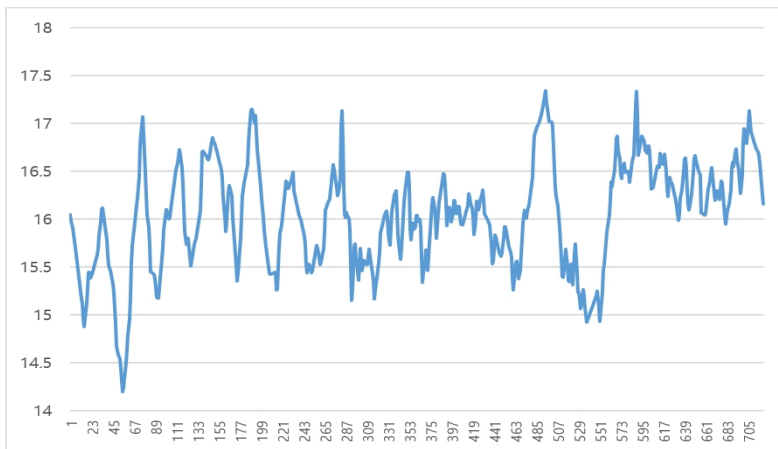


그림 18: 학습에 사용된 SO_x 배출 농도 60초 데이터

4.1.1 NARX 모델 최적화

1) Input Delay 변화 실험

NARX 모델에 다른 실험 조건(Hidden Unit : 10, Training Function : Trainlm)은 동일하게 하고 Input Delay를 2 ~ 50까지 변화시켜 예측 성능을 비교하였다. 그림 19의 실험결과에 따르면 Input Delay = 10일 때 예측 성능이 가장 좋으나 10 이하의 Data를 참조하는 경우는 예측 성능이 모두 좋으며 40이 넘어가면 예측 성능이 크게 감소한다.

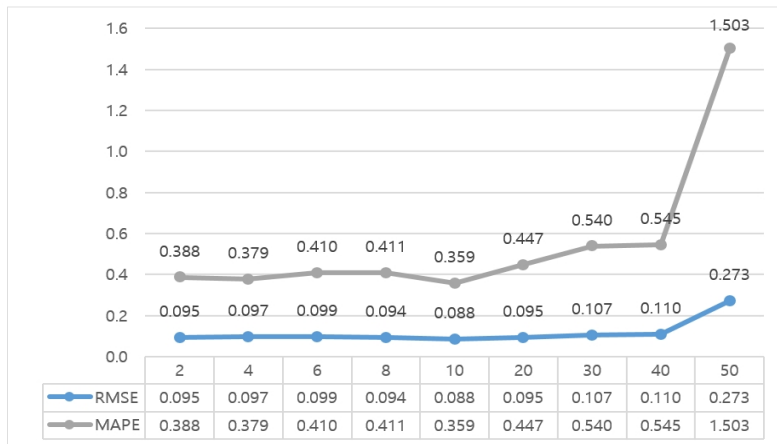


그림 19: NARX Model 60s Data Input Delay 실험 결과

2) Hidden Unit

다른 실험 조건(Input Delay : 10, Training Function : Trainlm)은 동일하게 고정한 상태에서 Hidden Unit을 2 ~ 40까지 변화시켜 예측 성능을 비교하였다. 그림 20의 실험결과에 따르면 Hidden Unit = 6일 때 예측 성능이 가장 좋고 2 ~ 8로 설정한 경우 예측 성능이 모두 좋으며 20보다 Unit 개수가 많아지면 예측 성능이 감소한다.

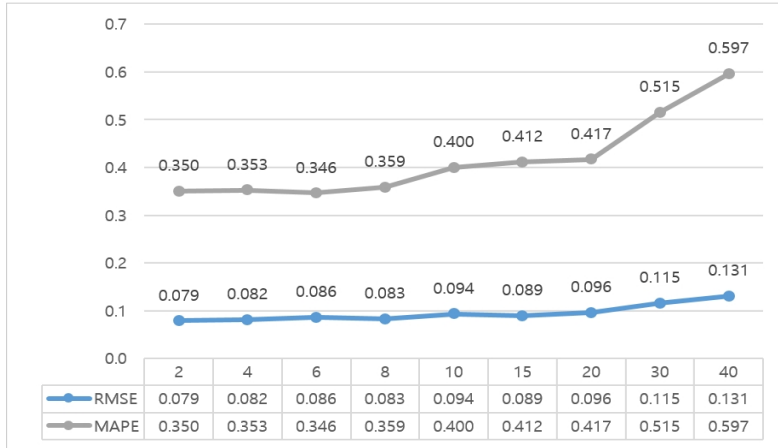


그림 20: NARX Model 60s Data Hidden Unit 실험 결과

3) Training Function

다른 실험 조건(Input Delay : 10, Hidden Unit : 6)은 이전 실험의 결과를 반영하고 Training Function(trainlm / trainrp / traingdx)을 변화하여 실험하였다. 그림 21의 결과와 같이 trainlm에서 예측 성능이 가장 좋음을 알 수 있다.

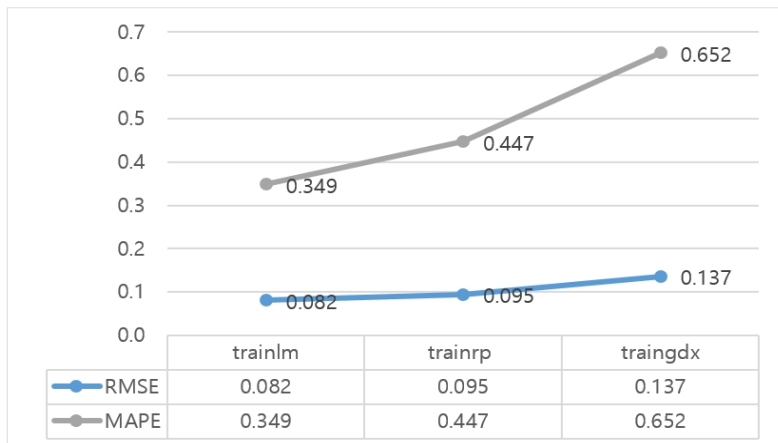


그림 21: NARX Model 60s Data Training Function 실험 결과

4.1.2 LSTM 모델 최적화

1) Training Function

LSTM 모델에 다른 실험 조건(Hidden Layer : 10, Learning Rate : 0.005, Epochs : 300, Drop : 75(0.2))은 유지 하고 Training Function(Adam, Rmsprop ,SGDM)을 변화하여 실험하였다. 그림 22의 결과와 같이 Adam 에서 예측 성능이 가장 좋다는 것을 알 수 있다.

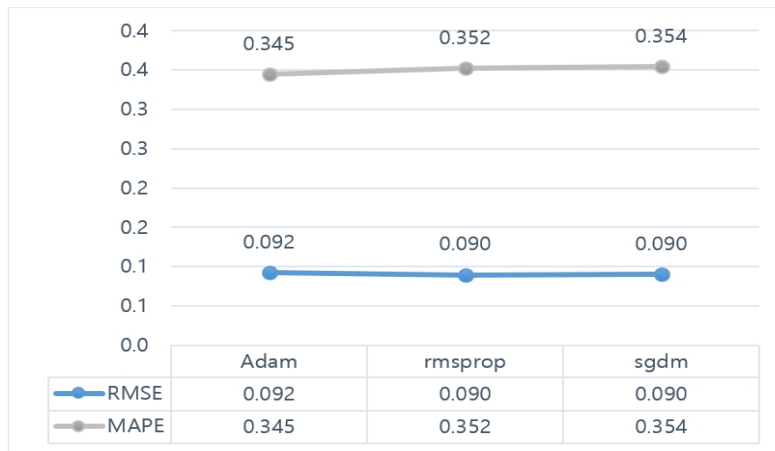


그림 22: LSTM Model 60s Data Training Function 실험 결과

2) Hidden Layer

Training Function은 앞선 실험을 반영하여 Adam을 적용하고 실험 조건(Learning Rate : 0.005, Epochs : 300, Drop : 75(0.2), Training Function : Adam)을 고정한 후 Hidden Layer를 3 ~ 300까지 변화하여 실험하였다. 그림 23의 결과와 같이 Hidden Layer = 8에서 예측 성능이 가장 좋으며 8~100까지는 예측 성능이 좋고 300에서 예측 성능이 저하된다.

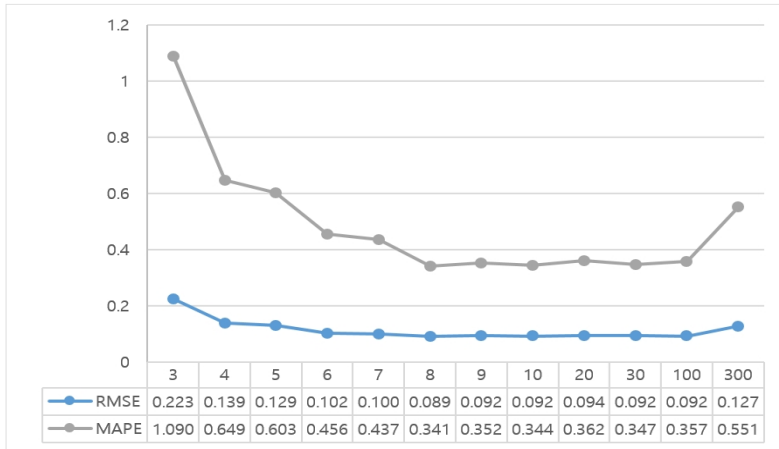


그림 23: LSTM Model 60s Data Hidden Layer 실험 결과

3) Learning Rate

LSTM 모델에 실험 조건(Hidden Layer : 8, Epochs : 300, Drop : 75(0.2), Training Function : Adam)은 앞선 실험을 반영하고 Learning Rate를 0.005 ~ 0.03까지 변화하여 실험한다. 그림 24의 결과와 같이 0.025에서 예측 성능이 가장 좋다.

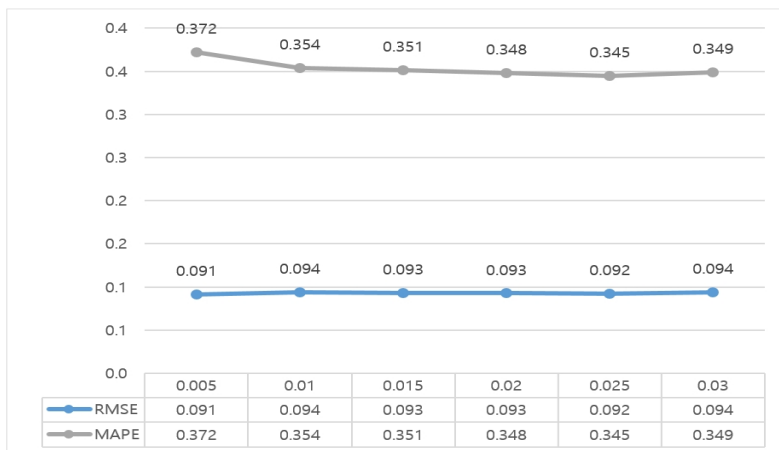


그림 24: LSTM Model 60s Data Learning Rate 실험 결과

4) Epochs

LSTM 모델에 다른 실험 조건(Hidden Layer : 8, Learning Rate : 0.025, Training Function : Adam)은 고정 하고 Epochs를 80~400까지 변화하여 실험하였다. 그림 25의 결과와 같이 Epochs 200에서 가장 예측 성능이 좋고 100 ~ 300사이에서 좋은 성능을 보인다. 괄호 안의 숫자는 Drop을 나타낸 것으로 200 Epochs의 경우 75 Epochs마다 Learning Rate값에 0.2를 곱하여 점점 감소하도록 적응적 학습률 기법을 반영한 것이다.

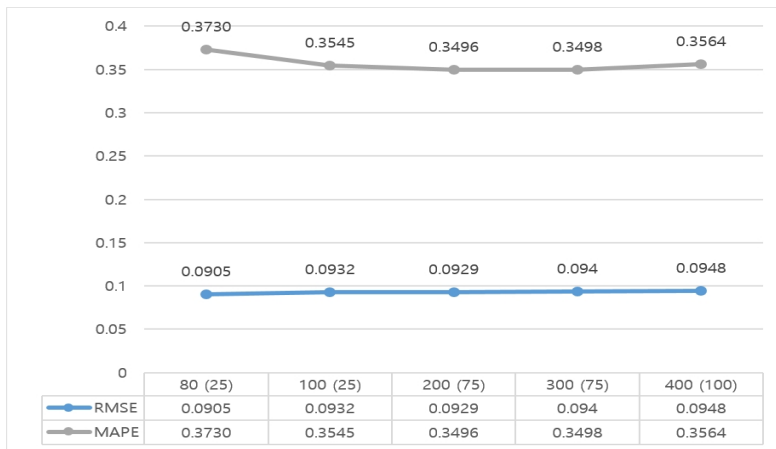


그림 25: LSTM Model 60s Data Epochs 실험 결과

4.1.3 60초 Event 데이터 검증

최적화된 모델을 검증하기 위해 60초 샘플링 간격으로 탈황공정의 Event Data(2017.08.26, 12:00~20:00)를 취득하였다. 검증에 사용된 SOx 배출 농도 데이터는 그림 26와 같다. 학습에 사용된 데이터 보다 최대값과 최저값의 범위가 크고 순간 변화가 큰 데이터이다.

본 연구의 목적이 이와 같은 변화를 예측할 수 있는 모델을 제안하는 것이므로 Event 데이터를 예측 모델의 검증에 사용하였다. 재학습 데이터(2017.08.26, 00:00~12:00)는 Event 데이터보다 앞선 시점의 12시간 연속 데이터로 학습 데이터가 예측성능에 미치는 영향을 비교하기 위해 사용되었다.

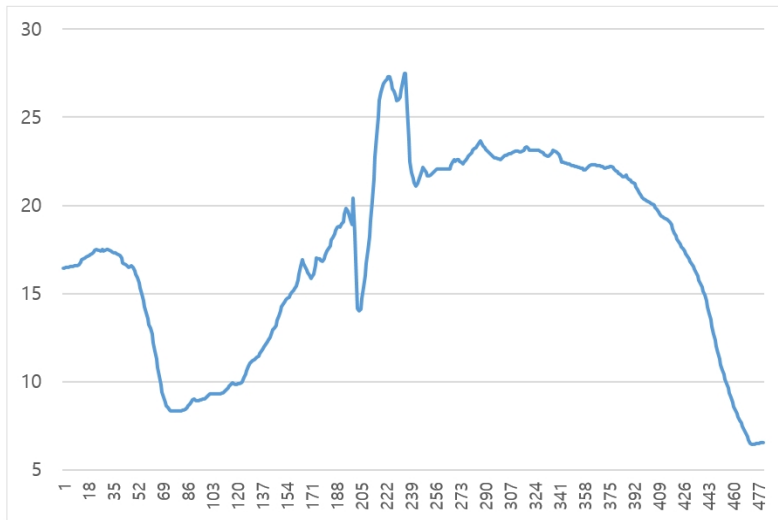


그림 26: 검증에 사용된 SOx 배출 농도 60초 Event 데이터

1) NARX 모델 검증

최적화된 하이퍼 파라미터(Input Delay : 10, Hidden Unit : 6, Training Function : trainlm)로 설정한 NARX 모델을 학습시킨 후 Event Data를 예측하여 측정값과 비교하는 실험을 하였다. 그림 27의 결과를 보면 NARX 최적화 모델은 Event Data의 예측 성능이 좋지 않다. 동일한 모델을 Event Data 이전 시점의 데이터를 취득하여 재학습시킨 후 검증 하였고 그림 28의 결과를 보면 학습 Data에 의해 성능이 향상되었다.

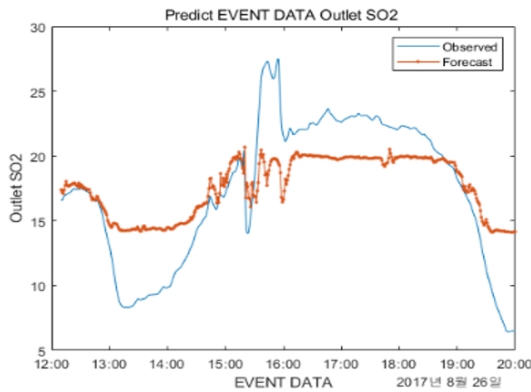


그림 27: NARX 60s Data 최적화 모델 Event Data 예측

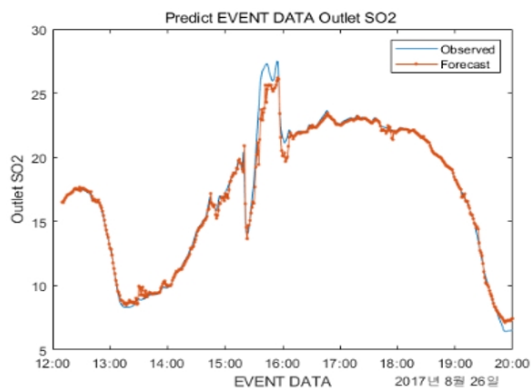


그림 28: NARX 60s Data 최적화 모델 재학습 후 Event Data 예측

하이퍼 파라미터와 학습 데이터 중 예측 성능에 영향을 미치는 요인을 확인하기 위해 하이퍼 파라미터를 가장 작은 값으로 변경한 NARX 최소화 모델(Input Delay : 2, Hidden Unit : 2, Training Function : trainlm)을 만들어 Event Data를 예측하였다.

그림 29의 결과를 보면 NARX 최소화 모델은 최적화 모델과 유사하게 예측 성능은 좋지 않다. 최소화 모델을 Event Data와 동일한 날의 이전 시점 데이터를 취득하여 재학습시킨 후 예측 한 결과 그림 30과 같이 예측 성능이 향상되었다. 이 실험을 통해 NARX 모델은 하이퍼 파라미터 보다 학습 데이터의 의존도가 높음을 알 수 있다.

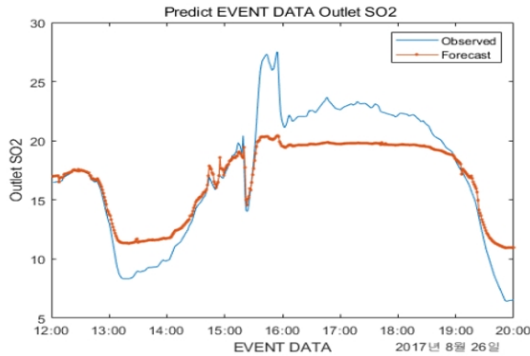


그림 29: NARX 60s Data 최소화 모델 Event Data 예측

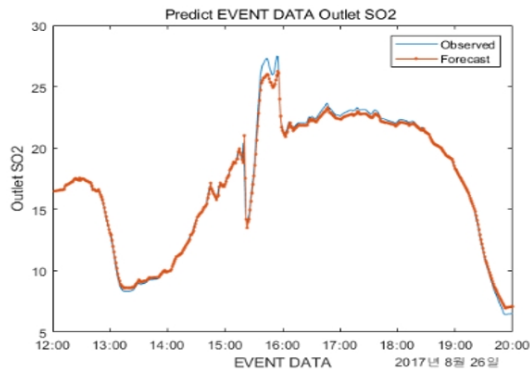


그림 30: NARX 60s Data 최소화 모델 재학습 후 Event Data 예측

최적화된 LSTM 모델(Hidden Layer : 8, Learning Rate : 0.025, Training Function : Adam, Epochs : 200, Drop : 75(0.2))을 학습 Data로 학습시킨 후 Event Data를 예측하여 측정값과 비교하였다. 그림 31의 결과를 보면 LSTM 최적화 모델은 Event Data의 예측 성능이 좋다는 것을 알 수 있다. 동일한 모델을 Event Data와 같은 날의 이전 시점 데이터를 취득하여 재학습 한 후 실험한 결과 예측 성능이 감소되었다. 그림 31의 결과와 32의 결과를 비교하면 학습 데이터로 최적화한 모델이 예측 성능이 더 좋다는 것을 알 수 있다.

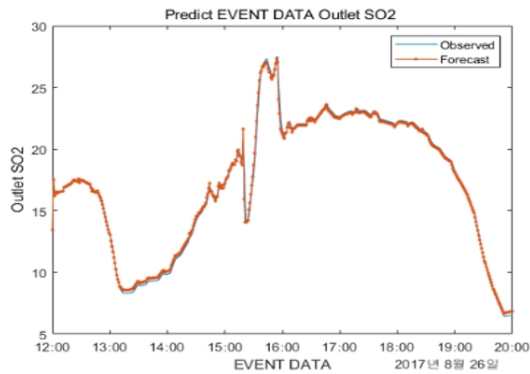


그림 31: LSTM 60s Data 최적화 모델 Event Data 예측

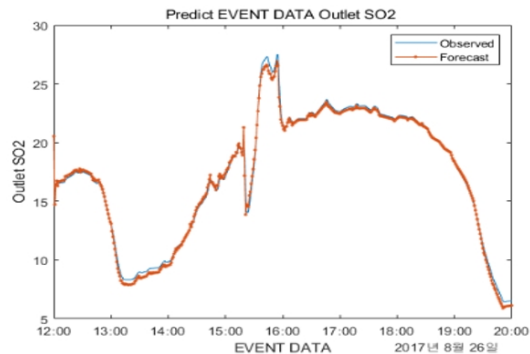


그림 32: LSTM 60s Data 최적화 모델 재학습 후 Event Data 예측

최적화된 NARX 모델과 LSTM 모델의 예측 성능을 비교한 결과 Event Data 예측 성능은 RMSE = 0.295, MAPE = 1.296으로 LSTM 최적화 모델이 가장 좋은 결과를 보였다. LSTM 모델의 경우 학습 Data로 최적화한 모델이 가장 뛰어난 예측 성능을 보였고 NARX 모델은 재학습한 모델이 좋은 성능을 보였으며 하이퍼 파라미터 설정보다 학습 데이터에 대한 의존도가 더 높은 것으로 나타났다.

그림 33의 결과는 앞선 예측 실험의 결과를 표로 나타낸 것으로 그림 27 ~ 32의 결과를 순서대로 표현하였다. NARX 최소(TD)와 NARX 최소(ED)는 그림 29와 30의 실험 결과를 나타낸 것으로 최소화 모델을 학습 데이터와 재학습 데이터로 각각 학습시킨 후 예측성능을 비교한 결과이다. NARX 모델의 경우 최적화 모델과 최소화 모델 모두 재학습 시킨 결과가 예측성능이 더 좋다는 것을 알 수 있다.

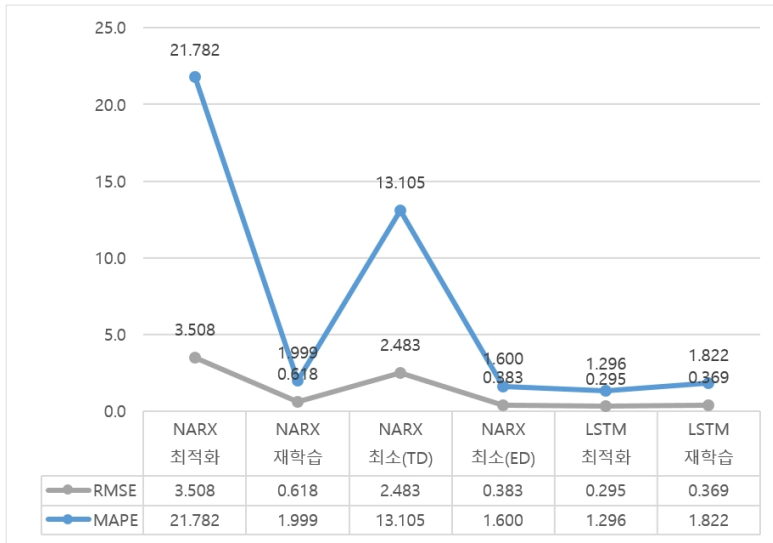


그림 33: 60s Data 최적화 모델 Event Data 검증 결과

4.2 120초 데이터 탈황예측 실험

모델을 최적화하기 위해 120초 샘플링 간격으로 황산화물(SO_x) 배출 농도 값과 Limestone Slurry valve 개도 값을 취득하였다. 학습에 사용된 SO_x 배출 농도 데이터는 그림 34와 같다. 그림 18과 동일한 기간의 데이터(2020.08.27, 00:00~12:00)를 취득하였지만 샘플링 간격이 2배이므로 학습 데이터의 개수가 반으로 줄어들게 되며 이는 Event 데이터(2017.08.26, 12:00~20:00)와 재학습 데이터(2017.08.26, 00:00~12:00)에도 동일하게 적용되었다.

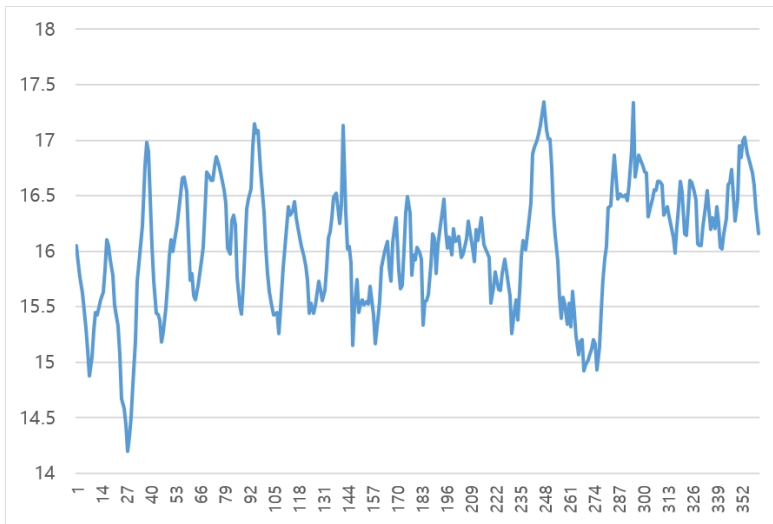


그림 34: 학습에 사용된 SO_x 배출 농도 120초 데이터

4.2.1 NARX 모델 최적화

1) Input Delay

NARX 모델에 다른 실험 조건(Hidden Unit : 10, Training Function : Trainlm)은 동일하게 하고 Input Delay를 2~30까지 변화시켜 예측 성능을 비교하였다. 그림 35의 실험결과에 따르면 Input Delay = 5 일 때 성능이 가장 좋고 20 이하의 Data를 참조하는 경우는 예측 성능이 모두 좋으며 30에서 예측 성능이 크게 감소한다.

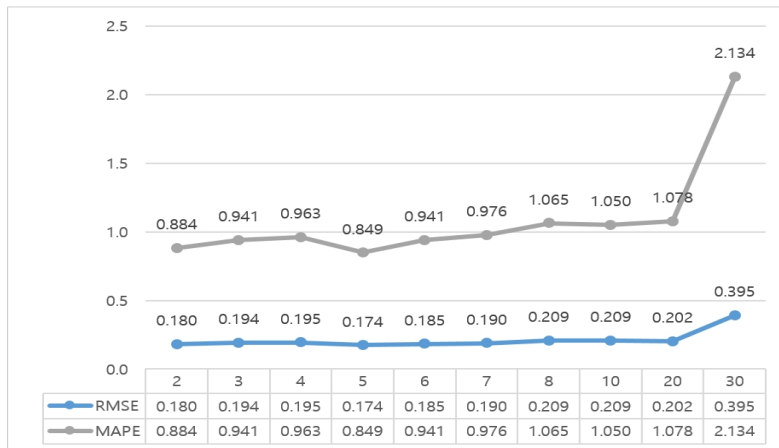


그림 35: NARX Model 120s Data Input Delay 실험 결과

2) Hidden Unit

앞의 실험결과를 반영한 실험 조건(Input Delay : 5, Training Function : Trainlm)에서 Hidden Unit을 2 ~ 40까지 변화하였다. 그림 36의 실험결과에 따르면 Hidden Unit = 8 일때 성능이 가장 좋으며 10보다 Unit 개수가 많아지면 MAPE 값이 점진적으로 상승하여 성능이 저하된다.

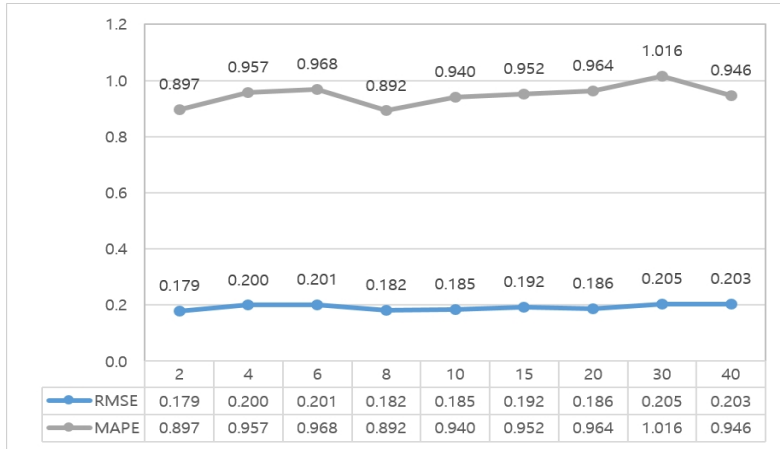


그림 36: NARX Model 120s Data Hidden Unit 실험 결과

3) Training Function

다른 실험 조건(Input Delay : 5, Hidden Unit : 8)은 앞의 실험 결과를 반영하여 고정하고 Training Function(trainlm / trainrp / traingdx)을 변화하여 실험하였다. 그림 37의 결과와 같이 trainlm에서 예측 성능이 가장 좋음을 알 수 있다.

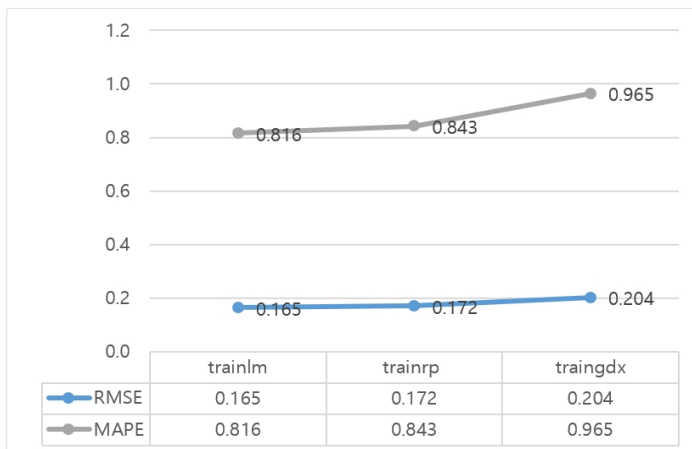


그림 37: NARX Model 120s Data Hidden Unit 실험 결과

4.2.2 LSTM 모델 최적화

1) Training Function

LSTM 모델에 실험 조건(Hidden Layer : 10, Learning Rate : 0.005, Epochs : 300, Drop : 75(0.2))은 LSTM의 기본 설정으로 고정하고 Training Function(Adam, Rmsprop, SGDM)만 변화하여 실험하였다. 그림 38의 결과와 같이 Adam에서 예측 성능이 가장 좋았다.

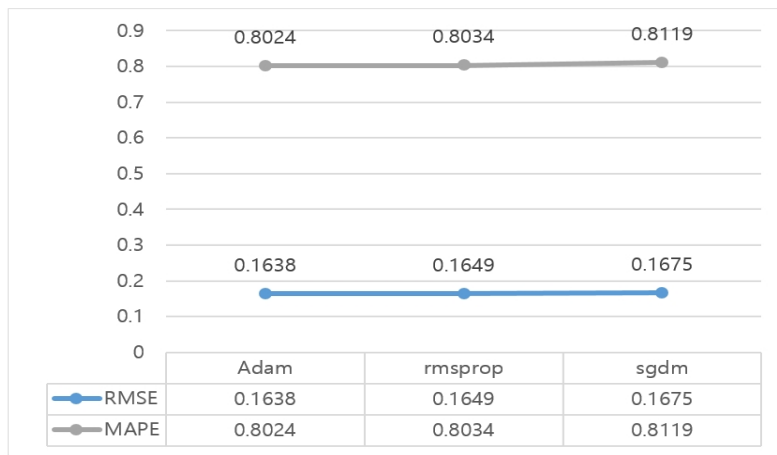


그림 38: LSTM Model 120s Data Training Function 실험 결과

2) Hidden Layer

LSTM 모델의 실험 조건(Learning Rate : 0.005, Epochs : 300, Drop : 75(0.2), Training Function : Adam)은 앞선 실험 결과를 반영하고 Hidden Layer를 1 ~ 40까지 변화하여 실험하였다. 그림 39의 결과를 보면 Hidden Layer = 10 에서 예측 성능이 가장 좋으나 3 ~ 30까지는 예측 성능이 좋고 40에서 예측 성능이 크게 저하된다.

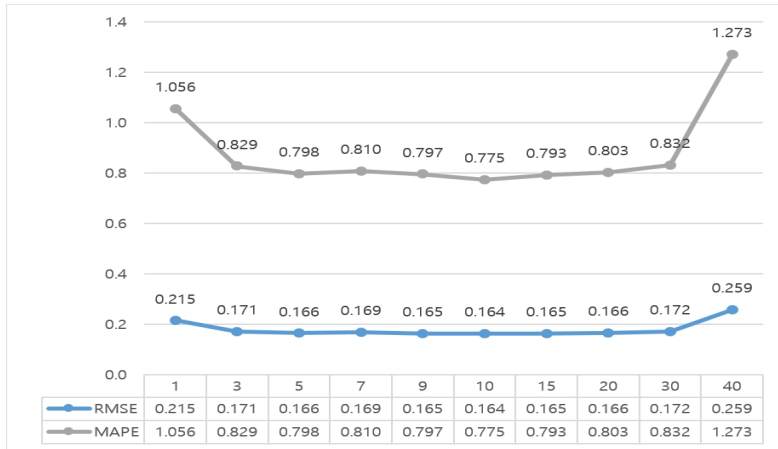


그림 39: LSTM Model 120s Data Hidden Layer 실험 결과

3) Learning Rate

LSTM 모델에 이전 실험 결과를 반영하여 다른 하이퍼 파라미터 (Hidden Layer : 10, Epochs : 300, Drop : 75(0.2), Training Function : Adam)는 고정 하고 Learning Rate를 0.005 ~ 0.035까지 변화하여 실험하였다. 그림 40의 결과와 같이 Learning Rate = 0.015 에서 예측 성능이 가장 좋다.

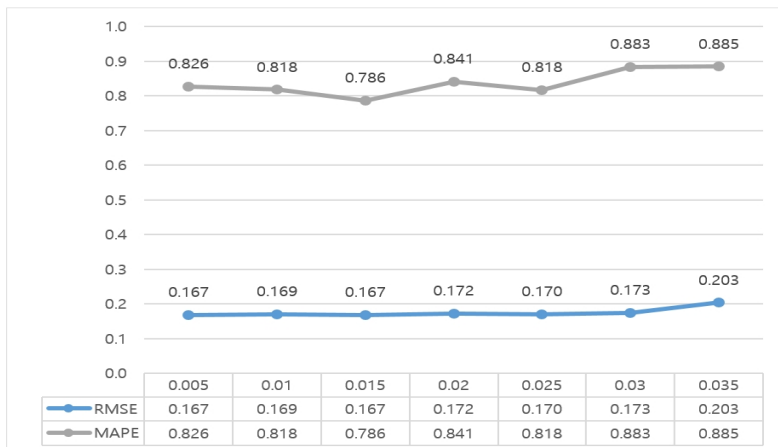


그림 40: LSTM Model 120s Data Learning Rate 실험 결과

4) Epochs

LSTM 모델에 이전 실험 결과를 반영하여 하이퍼 파라미터(Hidden Layer : 10, Learning Rate : 0.015, Training Function : Adam)를 설정하고 Epochs를 50 ~ 400까지 변화하여 실험하였다. 괄호 안의 숫자는 Drop를 나타낸 것으로 300 Epochs의 경우 100 Epochs마다 Learning Rate값에 0.2를 곱하여 Drop 하도록 설정한 것이다. 그림 41의 결과와 같이 300 Epochs에서 예측 성능이 가장 뛰어난 것을 확인할 수 있다.

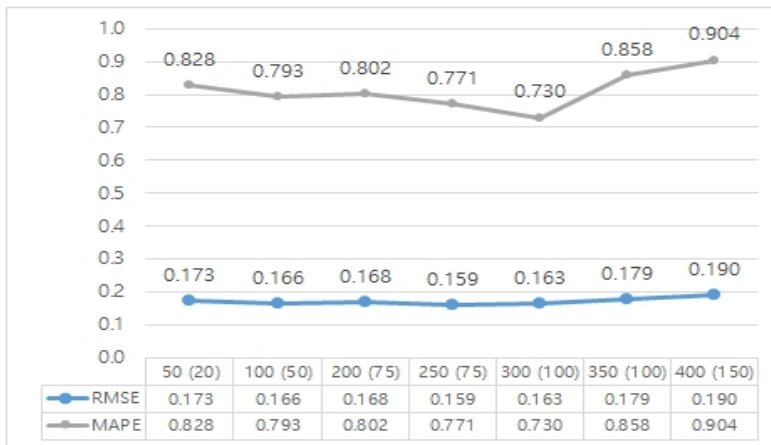


그림 41: LSTM Model 120s Data Epochs 실험 결과

4.2.3 120초 Event 데이터 검증

최적화 모델을 검증하기 위해 120초 샘플링 간격으로 탈황공정의 Event Data를 취득하였다. 검증에 사용된 SOx 배출 농도 데이터를 취득한 기간(2017.08.26, 12:00~20:00)은 앞서 설명한 그림 26과 같으며 샘플링 간격이 60초에서 120초로 두배로 길어지므로 데이터의 개수가 60초 데이터 보다 절반으로 줄어드는 것도 앞서 120초 학습 데이터(2020.08.27, 00:00~12:00)와 동일하다. Event 데이터보다 앞선 시점의 12시간 연속 데이터(2017.08.26, 00:00~12:00)를 취득하여 동일하게 재학습 데이터로 사용하였다.

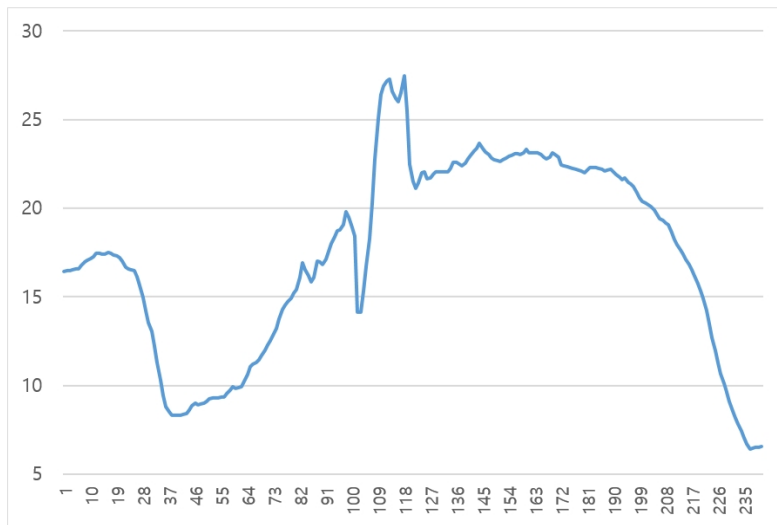


그림 42: 검증에 사용된 SOx 배출 농도 120초 Event 데이터

1) NARX 모델 검증

NARX 모델을 최적화된 하이퍼 파라미터(Input Delay : 5, Hidden Unit : 8, Training Function : trainlm)로 설정하고 학습시킨 후 Event Data를 예측하는 실험을 하였다.

그림 43의 결과를 보면 NARX 최적화 모델은 Event Data의 예측 성능이 좋지 않다. 동일한 모델을 Event Data 이전 시점의 데이터를 취득하여 재학습시킨 후 실험한 그림 44의 결과를 보면 예측 성능이 향상되었다.

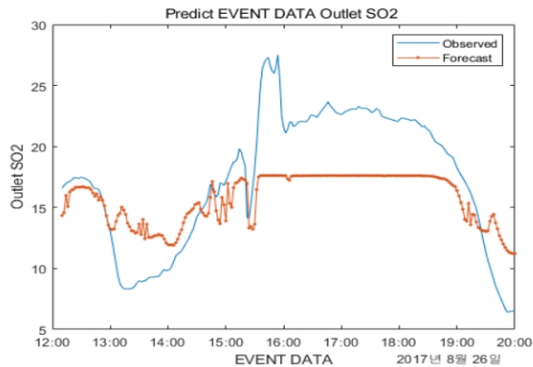


그림 43: NARX 120s Data 최적화 모델 Event Data 예측

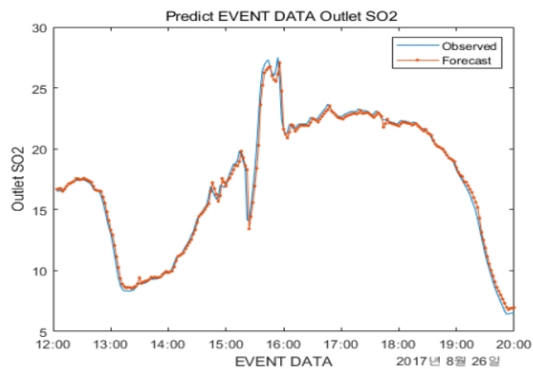


그림 44: NARX 120s Data 최적화 모델 Event Data 예측

앞서 60초 Event Data 검증 실험과 동일하게 하이퍼 파라미터와 학습 데이터 중 예측 성능에 영향을 미치는 요인을 확인하기 위해 하이퍼 파라미터를 가장 작은 값으로 변경한 NARX 최소화 모델(Input Delay : 2, Hidden Unit : 2, Training Function : trainlm)을 120초 Event Data에 적용하여 실험 하였다.

그림 45의 결과를 보면 NARX 최소화 모델은 최적화 모델과 유사하게 예측 성능이 좋지 않다. 최소화 모델을 Event Data와 동일한 날의 이전 시점 데이터를 취득하여 재학습시킨 후 실험 한 결과 그림 46와 같이 예측 성능이 향상되었다.

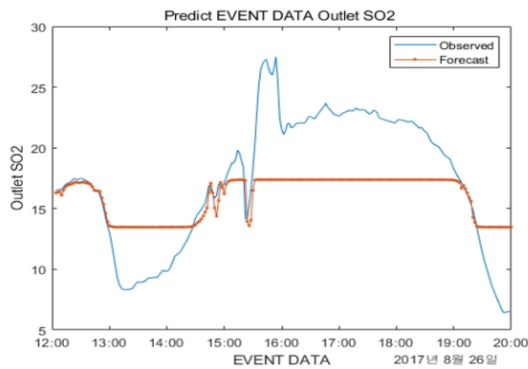


그림 45: NARX 120s Data 최소화 모델 Event Data 예측

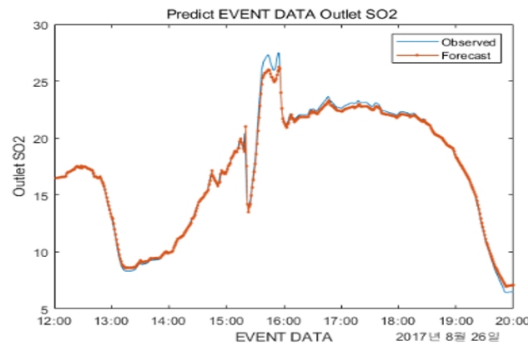


그림 46: NARX 120s Data 최소화 모델 재학습 후 Event Data 예측

최적화된 LSTM 모델(Hidden Layer : 10, Learning Rate : 0.015, Training Function : Adam, Epochs : 300, Drop : 100)을 학습 Data로 학습 시킨 후 Event Data를 예측하는 실험을 하였다.

그림 47의 결과를 보면 LSTM 120초 데이터 최적화 모델은 Event Data의 예측 성능이 좋다는 것을 알 수 있다. 동일하게 설정한 모델을 Event Data와 같은 날의 이전 시점 데이터를 취득하여 재학습시킨 후 실험을 하였다. 그림 48의 결과를 보면 학습 데이터로 최적화 한 모델이 Event Data로 재학습한 모델보다 예측 성능이 더 좋다는 것을 알 수 있다.

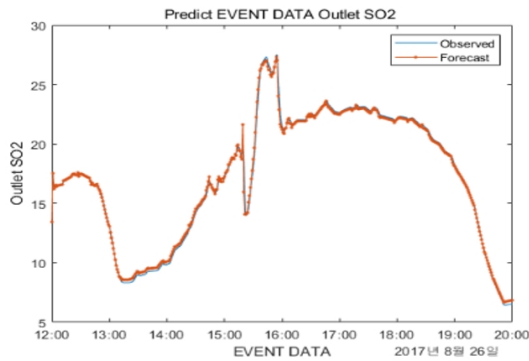


그림 47: LSTM 120s Data 최적화 모델 Event Data 예측

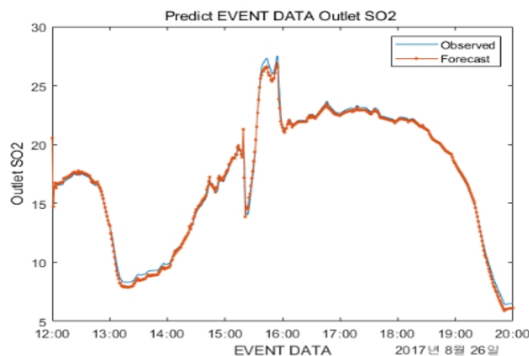


그림 48: LSTM 120s Data 최적화 모델 재학습 후 Event Data 예측

최적화된 NARX 모델과 LSTM 모델의 예측 성능을 비교한 결과 Event Data 예측 성능은 LSTM 모델이 RMSE = 0.664, MAPE = 3.931로 더 뛰어났다. NARX 모델의 경우에는 최소화 모델을 재학습 시켰을때 RMSE = 0.540, MAPE = 2.224로 가장 좋은 예측 결과를 보였다. 그림 49의 결과는 앞선 예측 실험의 결과를 표로 나타낸 것으로 그림 43 ~ 48의 결과를 순서대로 표현하였다. NARX 최소(TD)와 NARX 최소(ED)는 그림 45와 46의 실험 결과를 나타낸 것으로 최소화 모델을 학습 데이터와 재학습 데이터로 각각 학습시킨 후 예측성능을 비교한 결과이다.

그림 49의 결과를 통해 LSTM 최적화 모델은 재학습 한 경우보다 학습 데이터로 최적화 한 모델이 더 좋은 예측 성능을 보였으며 NARX 모델의 경우 최적화 모델과 최소화 모델 모두 재학습 시킨 결과가 예측 성능이 더 좋다는 것을 확인할 수 있다.

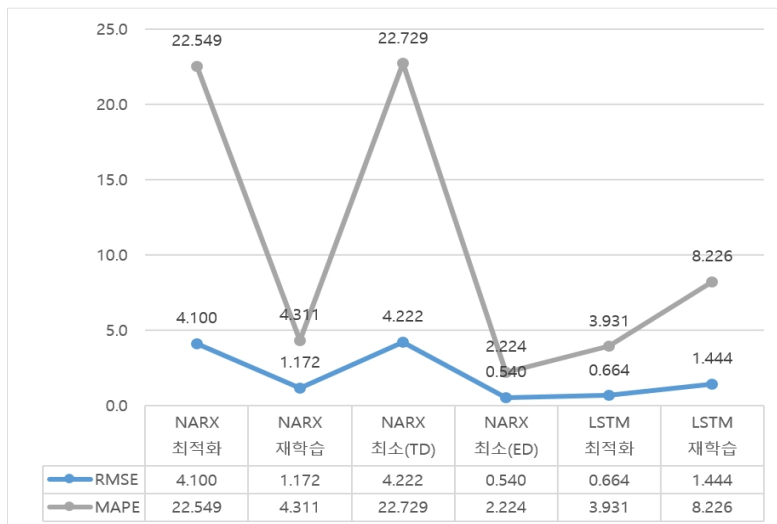


그림 49: 120s 최적화 모델 Event Data 검증 결과

제 5 장

탄종변화 실험 및 결과

국내 석탄화력 발전소는 유연탄을 수입하여 사용하며 다양한 석탄을 혼합하여 연소하기 때문에 연소 후 탈황공정으로 유입되는 황산화물의 농도는 실시간으로 변동된다. 석탄화력 발전소의 연료는 8시간동안 연소할 수 있는 용량의 사일로(Silo)에 저장된 후 미분기로 공급되므로 연료 변화 시점을 정확히 알기 어려워 탄종 변화를 포함할 수 있도록 6일(144시간)의 연속된 데이터(2021.01.11 00:00~01.17 00:00)를 취득하였다.

60초와 120초 샘플링 간격으로 황산화물(SO_x) 배출 농도 값과 Limestone Slurry valve 개도 값을 취득하였으며 실험에 사용된 데이터는 그림 50의 좌측 3일(72시간)의 데이터(2021.01.11 00:00~01.14 00:00)로 각 모델을 최적화 하고 우측 3일의 데이터(2021.01.14 00:00~01.17 00:00)로 예측 성능을 검증하는 실험을 하여 결론을 도출하였다.

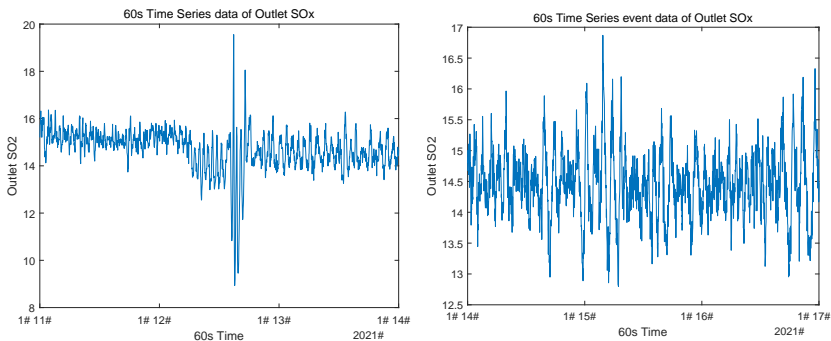


그림 50: 탄종변화 실험에 사용된 60초 데이터

표 4와 표 5의 자료를 통해 실험기간에 사용된 소비탄 분석 결과와 사용된 탄종을 알수 있다.

표 4: 탄종변화 실험 데이터 소비탄 분석 결과

실험일자	혼소 비율	소비량(톤)	발열량(kcal/kg)	유황분(%)
2021.01.11	30:41:29	7,102	6,770	0.40
2021.01.12	16:4:6:29:45	6,815	6,873	0.38
2021.01.13	12:14:6:69	6,847	6,690	0.45
2021.01.14	4:6:62:28	7,247	6,655	0.45
2021.01.15	11:11:16:26:22:14	6,981	6,722	0.46
2021.01.16	17:17:25:41	7,079	6,484	0.35

표 4에서 석탄의 발열량은 건식 기준(DB, 수분을 100% 제거한 상태의 시료를 기준으로 분석한 결과)으로 표기하였다.

표 5는 실험 기간중 혼합하여 사용한 탄종명을 표기한 것으로 탄종명이 중복하여 표기된 경우는 석탄이 입하된 항차가 다른 경우이다. 표 4에서 혼합하여 연소된 탄종의 분석결과를 알수 있으므로 항차별 입하된 석탄의 분석자료는 별도 표기하지 않았다.

표 5: 탄종변화 실험 혼소 탄종 데이터

실험일자	혼소 탄종
2021.01.11	SUEK(ID)+SUEK(RU)+SAMSUNG(AU)
2021.01.12	KCH(ID)+SUEK(ID)+STX(AU)+SUEK(RU)+SAMSUNG(AU)
2021.01.13	MOOLARBEN+STX(AU)+STX(AU)+SAMSUNG(AU)
2021.01.14	MOOLARBEN+STX(AU)+SAMSUNG(AU)+GLENCORE(GS)
2021.01.15	STX(AU)+SUEK(ID)+SUEK(RU)+FLAME(AU)+SAMSUNG(AU)+GLENCORE(GS)
2021.01.16	STX(AU)+SUEK(ID)+SUEK(RU)+FLAME(AU)

5.1 60초 데이터 탄중변화 실험

LSTM과 NARX 모델을 최적화하기 위해 60초 샘플링 간격으로 황산화물(SO_x) 배출 농도 값과 Limestone Slurry valve 개도 값을 취득하였다. 학습에 사용된 SO_x 배출 농도 데이터는 그림 51과 같다.

운전 데이터 중에서 변동이 큰 구간을 선정하고 탄중 변화에 의한 변동을 반영하기 위해 3일(72시간)의 데이터(2021.01.11 00:00~01.14 00:00)를 취득하였으며 학습데이터의 갯수는 4321개로 Training Data Set : Test Data Set은 9 : 1의 비율로 적용하였다.

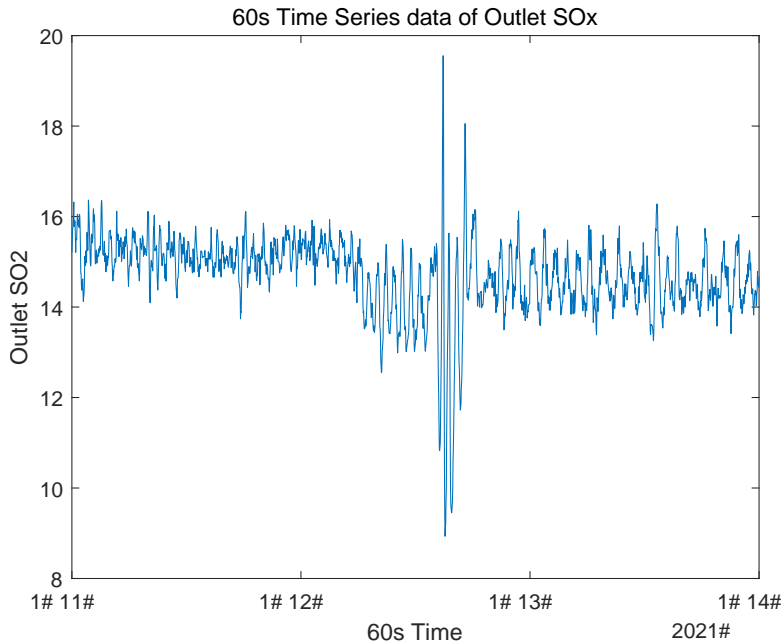


그림 51: 탄중변화 학습에 사용된 60초 데이터

5.1.1 NARX 모델 최적화

학습 데이터를 사용한 NARX 모델의 최적화 실험 결과를 표 6 ~ 8에 나타내었다. 실험조건을 표 하단에 정리하였는데 실험 할 하이퍼 파라미터를 제외하고 나머지 조건은 고정한 상태에서 표의 순서대로 실험을 하였으며 실험 결과를 반영하여 최적화 모델을 도출하였다.

표 6: 60s Data NARX Hyper Parameter : Training Function

Training Function	trainlm	trainrp	traingdx
RMSE	0.1021	0.1506	0.2321
MAPE	0.5072	0.8113	1.2295

[실험 조건] Training Function : change, Input Delay : 10, Hidden Unit : 10 으로 실험을 하였으며 trainlm이 가장 높은 예측 성능을 보였다.

표 7: 60s Data NARX Hyper Parameter : Input Delay

Input Delay	2	10	20	30	40	50
RMSE	0.1061	0.1066	0.1037	0.1023	0.1041	0.1081
MAPE	0.5273	0.5392	0.5124	0.4979	0.5147	0.5424

앞선 실험의 결과를 반영하여 [실험 조건] Training Function : trainlm, Input Delay : change, Hidden Unit : 10 으로 실험을 하였고 Input Delay = 30 에서 가장 높은 예측 성능을 보였다.

표 8: 60s Data NARX Hyper Parameter : Hidden Unit

Hidden Unit	2	4	5	10	20	30
RMSE	0.1088	0.1050	0.1029	0.1044	0.1047	0.1119
MAPE	0.5303	0.5131	0.5021	0.5182	0.5070	0.5650

앞선 실험의 결과를 반영하여 [실험 조건] Training Function : trainlm, Input Delay : 30, Hidden Unit : change 으로 실험을 하였고 Hidden Unit = 5 에서 가장 높은 예측 성능을 보였다.

5.1.2 LSTM 모델 최적화

LSTM 모델의 최적화 실험도 동일하게 수행하였다. 실험결과는 표 9 ~ 12에 나타내었으며 실험조건은 해당 하이퍼 파라미터를 제외하고 나머지 조건은 고정한 상태에서 학습시킨 후 성능평가를 하였다.

표 9: 60s Data LSTM Hyper Parameter : Training Function

Training Function	adam	rmsprop	sgdm
RMSE	0.1056	0.1045	0.1079
MAPE	0.5286	0.5245	0.5395

[실험 조건] Training Function : change, Hidden Layer : 10, Learning Rate : 0.025, Epochs : 200(Drop : 75, 0.2)으로 실험하였고 Training Function = rmsprop 에서 가장 좋은 예측 성능을 보였다.

표 10: 60s Data LSTM Hyper Parameter : Hidden Layer

Hidden Layer	1	5	10	50	100	200
RMSE	0.1071	0.1048	0.1048	0.1037	0.1041	0.1065
MAPE	0.5575	0.5266	0.5278	0.5133	0.5228	0.5295

앞선 실험의 결과를 반영하여 [실험 조건] Training Function : rmsprop, Hidden Layer : change, Learning Rate : 0.025, Epochs : 200 (Drop : 75, 0.2) 으로 실험하였고 Hidden Layer = 50에서 가장 높은 예측 성능을 보였다.

표 11: 60s Data LSTM Hyper Parameter : Learning Rate

Learning Rate	0.010	0.015	0.020	0.025	0.030	0.035
RMSE	0.1044	0.1041	0.1038	0.1041	0.1047	0.1045
MAPE	0.5293	0.5210	0.5185	0.5184	0.5250	0.5227

앞선 실험의 결과를 반영하여 [실험 조건] Training Function : rm-sprop, Hidden Layer : 50, Learning Rate : change, Epochs : 200 (Drop : 75, 0.2)로 실험하였으며 Learning Rate = 0.025에서 가장 높은 예측 성능을 보였다.

표 12: 60s Data LSTM Hyper Parameter : Epochs

Epochs	100	150	200	250	300	400
Drop(0.2)	50	75	75	75	75	75
RMSE	0.1079	0.1119	0.1042	0.1037	0.1048	0.1038
MAPE	0.5221	0.5787	0.5233	0.5142	0.5258	0.5187

앞선 실험의 결과를 반영하여 [실험 조건] Training Function : rm-sprop, Hidden Layer : 50, Learning Rate : 0.025, Epochs : change (Drop : change, 0.2)로 실험하였으며 Epochs = 250에서 가장 높은 예측 성능을 보였다.

Epochs의 값을 변화할때 적절한 값으로 Drop 주기를 변경해 주었으며 Drop = 75(0.2)의 의미는 75 Epochs 마다 Learning Rate 값에 0.2를 곱하여 학습 시작시에는 큰 값으로 학습률을 변화시키고 학습이 진행될수록 작은 값으로 학습률을 변화시키는 것으로 적응적 학습률 방식을 사용하였다.

5.1.3 60초 탄종변화 데이터 예측 검증

학습 데이터 이후 3일의 데이터(2021.01.14 00:00~01.17 00:00)를 취득하여 NARX와 LSTM 최적화 모델의 예측 성능을 검증하였다. 검증에 사용된 데이터는 그림 52와 같으며 데이터의 갯수는 4321개 이다.

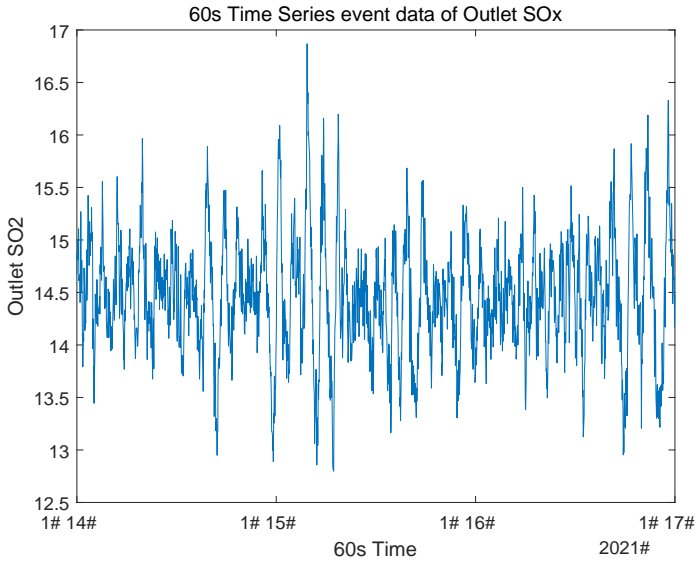


그림 52: 검증에 사용된 탄종변화 60초 데이터

그림 52의 데이터를 앞서 학습한 NARX와 LSTM 최적화 모델에 적용하여 예측을 시행하였고 예측결과를 성능검증 기준에 따라 RMSE와 MAPE로 나타내었다. 표 13에 의해 NARX 최적화 모델 보다 LSTM 최적화 모델이 예측 성능이 더 좋음을 알수 있다.

표 13: 60초 탄종변화 데이터 최적화 모델 검증 결과

평가기준	NARX 모델	LSTM 모델
RMES	0.1082	0.0705
MAPE	0.5568	0.3730

NARX 최적화 모델의 탄중변화 데이터 검증 결과를 그림 53에 나타내었다. NARX 최적화 모델의 하이퍼 파라미터는 다음 실험조건과 같다.

[실험 조건] Training Function : trainlm, Input Delay : 30, Hidden Unit = 5

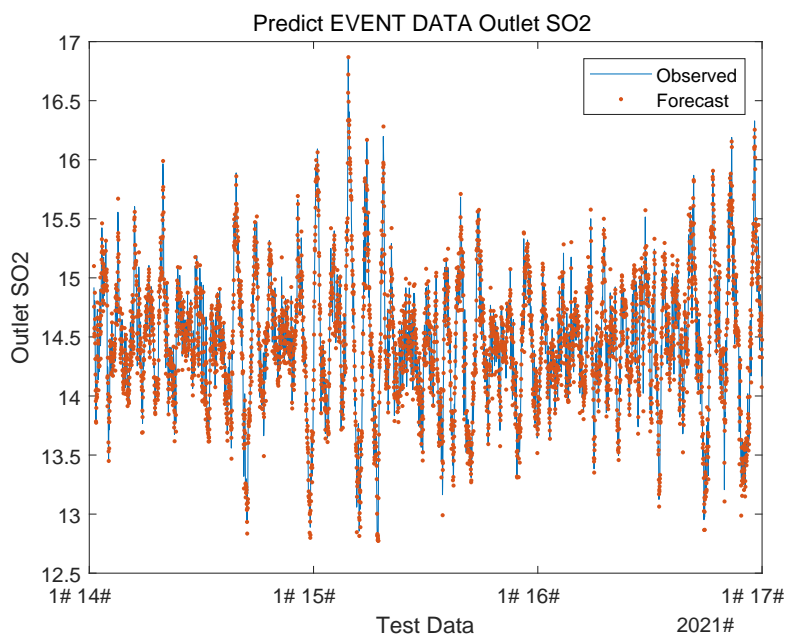


그림 53: 60초 탄중변화 데이터 NARX 모델 검증결과

LSTM 최적화 모델의 탄중변화 데이터 검증 결과를 그림 54에 나타내었다. LSTM 최적화 모델의 하이퍼 파라미터는 다음 실험조건과 같다.

[실험 조건] Training Function : rmsprop, Hidden Layer : 50,

Learning Rate : 0.025, Epochs : 250 (Drop : 75, 0.2)

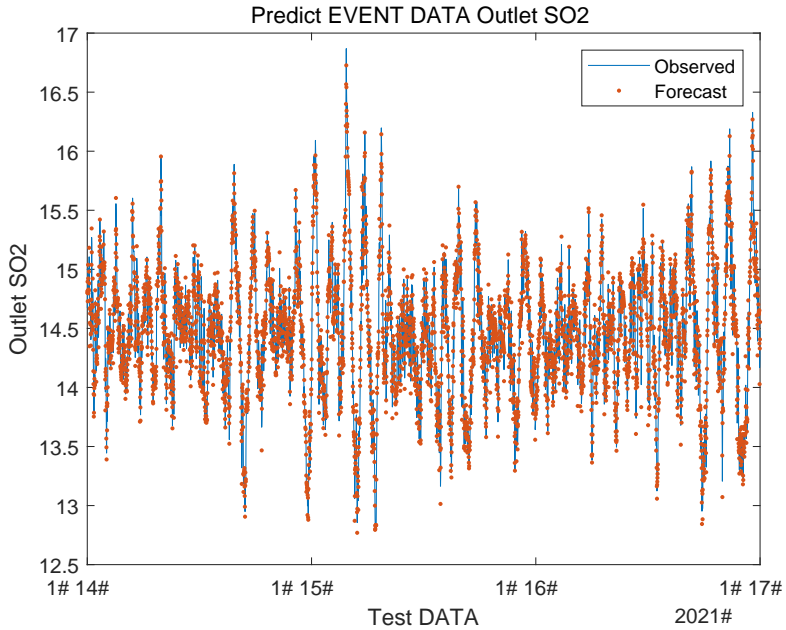


그림 54: 60초 탄중변화 데이터 LSTM 모델 검증결과

그림 53과 그림 54 에서 본 논문에서 제안한 최적화 모델이 탄중변화와 상관없이 황산화물(SOx) 배출 농도값을 잘 예측하고 있음을 확인할 수 있다.

5.2 120초 데이터 탄중변화 실험

LSTM과 NARX 모델을 최적화하기 위해 120초 샘플링 간격으로 황산화물(SOx) 배출 농도 값과 Limestone Slurry valve 개도 값을 취득하였고 학습에 사용된 SOx 배출 농도 데이터는 그림 55과 같다.

60초 학습 데이터와 동일한 기간의 데이터(2021.01.11 00:00~01.14 00:00)를 사용하였으나 샘플링 간격이 크므로 데이터의 갯수는 절반 수준인 2161개의 데이터를 사용하였다. 운전 데이터 중에서 변동이 큰 구간을 선정하고 탄중 변화에 의한 변동을 반영하기 위해 3일(72시간)의 데이터를 취득하였으며 Training Data Set : Test Data Set은 9 : 1의 비율로 적용하였다.

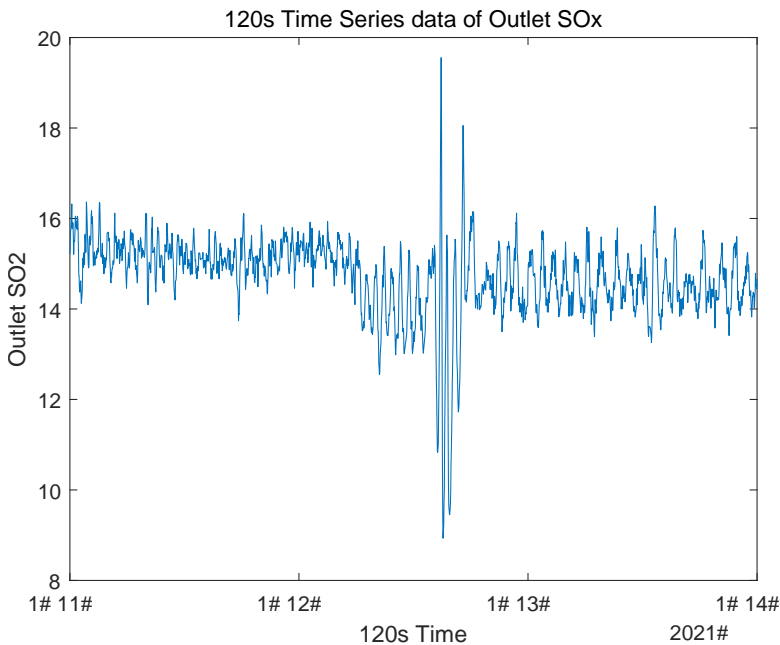


그림 55: 탄중변화 학습에 사용된 120초 데이터

5.2.1 NARX 모델 최적화

그림 55의 학습 데이터를 사용한 NARX 모델의 최적화 실험 결과를 표 14 ~ 16에 나타내었다. 각 하이퍼 파라미터의 실험조건은 표 하단에 명시되어 있으며 순서대로 실험을 하여 이전 하이퍼 파라미터의 결과를 적용하고 나머지 조건은 고정된 상태에서 성능평가를 하였다.

표 14: 120s Data NARX Hyper Parameter : Training Function

Training Function	trainlm	trainrp	traingdx
RMSE	0.1846	0.2074	0.2668
MAPE	1.0097	1.1292	1.3870

[실험 조건] Training Function : change, Input Delay : 10, Hidden Unit : 10
 으로 실험 결과 raining Function = trainlm 에서 예측 성능이 가장 우수하였다.

표 15: 120s Data NARX Hyper Parameter : Input Delay

Input Delay	2	5	8	10	20	30
RMSE	0.1866	0.1843	0.1828	0.1860	0.1864	0.2406
MAPE	1.0221	1.0096	1.0022	1.0119	1.0104	1.3322

[실험 조건] Training Function : trainlm, Input Delay : change, Hidden Unit : 10
 으로 실험 결과 Input Delay = 8 에서 예측 성능이 가장 우수하였다.

표 16: 120s Data NARX Hyper Parameter : Hidden Unit

Hidden Unit	2	4	6	8	10	20
RMSE	0.1875	0.1859	0.1866	0.1787	0.1827	0.1849
MAPE	1.0231	1.0138	1.0165	0.9709	0.9883	1.0100

[실험 조건] Training Function : trainlm, Input Delay : 8, Hidden Unit : change
 으로 실험 결과 Hidden Unit = 8 에서 예측 성능이 가장 우수하였으며 최적화 모델을 도출하였다.

5.2.2 LSTM 모델 최적화

그림 55의 학습 데이터를 사용한 LSTM 모델의 최적화 실험 결과를 표 17 ~ 표 20에 나타내었다. NARX 모델의 실험과 동일하게 각 하이퍼 파라미터의 실험조건은 이전 하이퍼 파라미터의 결과를 적용하고 나머지 조건은 고정한 상태에서 성능평가를 하였다.

표 17: 120s Data LSTM Hyper Parameter : Training Function

Training Function	adam	rmsprop	sgdm
RMSE	0.1922	0.1851	0.1963
MAPE	1.0506	1.0158	1.0802

[실험 조건] Training Function : change, Hidden Layer : 10,

Learning Rate : 0.025, Epochs : 200 (Drop : 75, 0.2)

실험 결과 Training Function = rmsprop 에서 가장 예측 성능이 우수하였다.

표 18: 120s Data LSTM Hyper Parameter : Hidden Layer

Hidden Layer	1	5	10	30	50	100
RMSE	0.1991	0.1989	0.1851	0.1878	0.1906	0.2098
MAPE	1.1036	1.0836	1.0107	1.0288	1.0504	1.0986

[실험 조건] Training Function : rmsprop, Hidden Layer : change,

Learning Rate : 0.025, Epochs : 200 (Drop : 75, 0.2)

실험 결과 Hidden Layer = 10 에서 가장 예측 성능이 우수하였다.

60초 데이터의 실험에서는 Hidden Layer = 50에서 최적화 되었는데 120초 데이터에서는 낮은 값에서 더 좋은 예측성능을 보였다.

표 19: 120s Data LSTM Hyper Parameter : Learning Rate

Learning Rate	0.020	0.025	0.030	0.035	0.040	0.045
RMSE	0.1867	0.1887	0.1903	0.1856	0.1835	0.1883
MAPE	1.0326	1.0431	1.0490	1.0171	1.0046	1.0401

[실험 조건] Training Function : rmsprop, Hidden Layer : 10,

Learning Rate : change, Epochs : 200 (Drop : 75, 0.2)

실험 결과 Learning Rate = 0.040에서 가장 예측 성능이 우수하였다. 60초 데이터 실험에서는 Learning Rate = 0.025에서 최적화 되었는데 120초 데이터에서는 높은 값에서 더 좋은 예측성능을 보였다.

표 20: 120s Data LSTM Hyper Parameter : Epochs

Epochs	100	150	200	250	300	400
Drop(0.2)	50	75	75	75	75	75
RMSE	0.1979	0.1924	0.1869	0.1879	0.1884	0.1901
MAPE	1.0898	1.0557	1.0267	1.0315	1.0332	1.0447

[실험 조건] Training Function : rmsprop, Hidden Layer : 10,

Learning Rate : 0.040, Epochs : change (Drop : change, 0.2)

실험 결과 Epochs = 200 값에서 가장 예측 성능이 우수하였다. Epochs 의 값을 변화할때 Drop 의 값도 같이 변화하여 실험 하였으며 Drop은 지정된 Epochs마다 Learning Rate값에 0.2를 곱하여 점진적으로 감소 하도록 설정한 것으로 적응적 학습률을 사용하였다.

5.2.3 120초 탄종변화 데이터 예측 검증

학습 데이터 이후 연속된 3일의 데이터(2021.01.14 00:00~01.17 00:00)를 취득하여 NARX와 LSTM 최적화 모델의 예측 성능을 검증하였다. 탄종변화 데이터는 그림 56에 나타내었으며 검증에 사용된 데이터의 갯수는 2161개 이다.

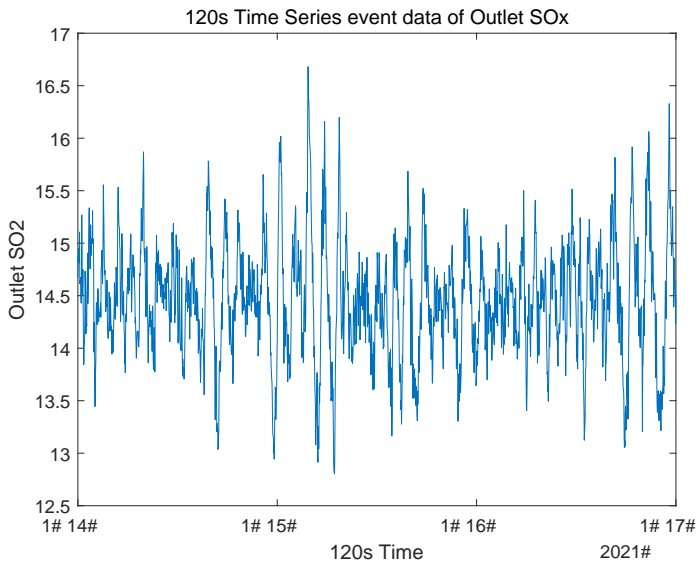


그림 56: 검증에 사용된 탄종변화 120초 데이터

그림 56의 데이터를 NARX와 LSTM 최적화 모델에 적용하여 예측을 시행하였고 예측결과를 표 21에 나타내었다. 실험 결과 NARX 최적화 모델 보다 LSTM 최적화 모델이 예측 성능이 더 좋음을 알수 있다.

표 21: 120초 탄종변화 데이터 최적화 모델 검증 결과

평가기준	NARX 모델	LSTM 모델
RMES	0.2279	0.1009
MAPE	1.1671	0.4689

NARX 최적화 모델의 탄중변화 데이터 검증 결과를 그림 57에 나타내었다. NARX 최적화 모델의 하이퍼 파라미터는 다음 실험조건과 같다.

[실험 조건] Training Function : trainlm, Input Delay : 8, Hidden Unit = 8

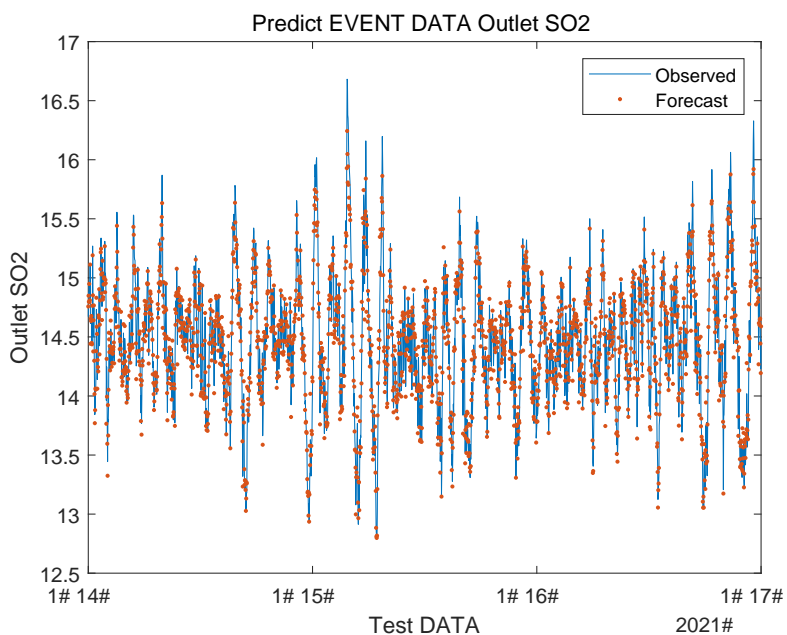


그림 57: 120초 탄중변화 데이터 NARX 모델 검증결과

LSTM 최적화 모델의 탄중변화 데이터 검증 결과를 그림 58에 나타내었다. LSTM 최적화 모델의 하이퍼 파라미터는 다음 실험조건과 같다.

[실험 조건] Training Function : rmsprop, Hidden Layer : 10,

Learning Rate : 0.040, Epochs : 200 (Drop : 75, 0.2)

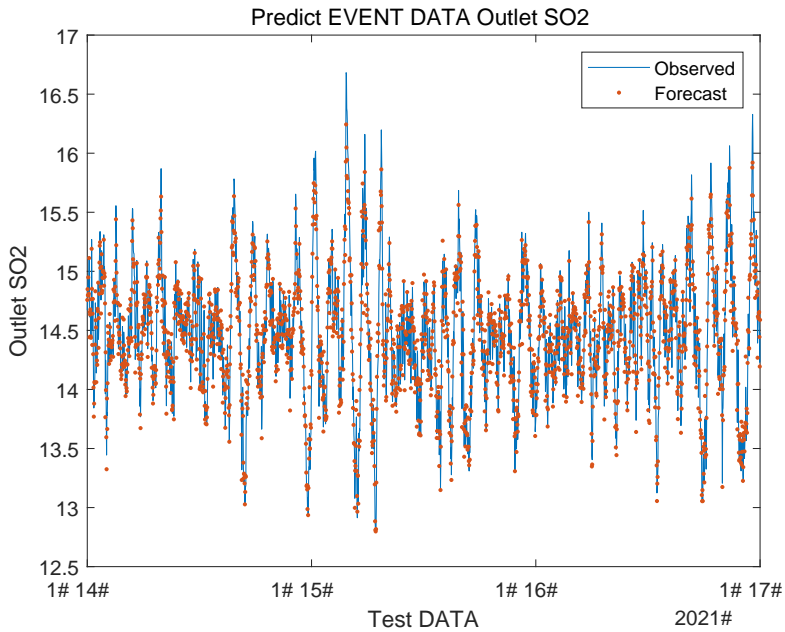


그림 58: 120초 탄중변화 데이터 LSTM 모델 검증결과

그림 57과 그림 58 에서 본 논문에서 제안한 최적화 모델이 탄중 변화에도 황산화물(SO_x) 배출 농도값을 잘 예측하고 있음을 확인할 수 있다. NARX 모델보다 LSTM 모델의 예측 성능이 더 우수하며 예측 결과 값이 측정 값의 변화를 더 잘 반영하고 있다.

제 6 장

결론

석탄화력 발전소는 다양한 계측기를 통해 시계열 데이터를 확보하고 공정 운전과 감시(Monitoring)에 활용하고 있으나 인공지능경망을 공정에 적용하는 연구는 활발하게 시도되지 않고 있다. 국내와는 달리 해외에서는 환경규제가 강화됨에 따라 석탄화력 발전소의 탈황공정과 탈질공정에 순환신경망 모델을 활용한 연구 사례가 있으며 국내에서도 다른 산업분야에서는 다양한 연구가 제안되고 있다.

정부의 기후변화 대응 정책에 따라 석탄화력 발전소에서 배출되는 대기오염물질에 대한 규제가 강화되고 있고 특히 황산화물의 경우 석탄 혼합 연소와 석회석 고갈에 따른 품질 저하로 배출 농도 예측을 통한 사전 대응이 필요한 상황이다. 따라서 본 논문은 석탄화력 발전소의 운전데이터를 활용하여 황산화물(SO_x) 배출 농도 예측 모델을 제안하고자 하였다.

시계열 데이터의 예측을 위해 인공지능경망(ANN)의 한 종류인 순환신경망(RNN)을 사용 하였고 순환신경망은 과거의 정보가 미래에 영향을 줄 수 있는 구조로 순차 데이터의 예측에 적합하나 데이터가 길어지면 장기 의존성 문제가 있다. 고유의 메모리 셀(Cell) 구조를 가지고 있어 이러한 문제를 해결한 LSTM(Long Short Term Memory network) 모델을 선정하였고 심층 신경망인 LSTM 모델과 비교하기 위해 단층 신경망인 NARX(Nonlinear Auto Regressive Neural Network with Exogenous Input) 모델을 사용하여 예측 실험을 설계 하였다.

3장에서 실험을 위해 탈황 공정에 영향을 미치는 변수를 도출하고 상관관계 분석을 통해 학습 데이터를 취득하기 위한 공정변수를 선별하였다. 공정의 시정수와 황산화물 농도 계측기의 측정 지연시간을 검토하여 60초와 120초 샘플링 간격을 선정하여 선별된 변수의 공정 운전 데이터를 취득하였다.

4장의 실험에서 학습 데이터는 60초와 120초 간격으로 취득된 두개의 데이터로 구성하였고 NARX와 LSTM의 모델에 각각 적용하여 예측 성능 실험을 하였으며 각 모델 별 학습데이터에 최적화된 하이퍼 파라미터를 도출하였다.

최적화 모델에 변동이 큰 Event 데이터를 적용하여 예측 성능을 검증하였다. 실험 결과 최적화 모델의 경우 60초와 120초 데이터 모두 LSTM 모델이 NARX 모델보다 예측 성능이 뛰어났으며 NARX 모델의 경우 최적화 모델이 Event 데이터의 예측에는 부적합한 결과를 나타내었다. Event 데이터와 같은 날 이전 시점에 취득한 데이터로 재학습 하여 실험한 결과 NARX 모델의 경우 예측 성능이 향상되었으나 LSTM 모델의 경우 학습 데이터에 최적화된 모델이 더 예측 성능이 뛰어났다.

NARX 모델의 경우 하이퍼 파라미터와 데이터에 의한 영향을 비교하기 위해 하이퍼 파라미터 설정을 최소화한 모델을 설정하고 동일한 실험을 한 결과 NARX 모델은 최적화 보다 데이터에 대한 의존성이 더 높은 것을 알 수 있었다.

LSTM 모델은 예측 성능이 좋으나 데이터 변화에 따라 모델의 최적화가 필요하고 NARX 모델은 좋은 성능을 내기 위해서 하이퍼 파라미터의 최적화도 중요하지만 학습 데이터를 적절하게 선정하는 것이 중요하다는 것을 확인하였다.

실험 결과를 종합하면 60초와 120초 데이터 모두 최적화 한 LSTM 모델이 NARX 모델보다 예측 성능이 좋았으며 재학습을 통해 NARX 모델의 예측성능이 증가하였으므로 예측하고자 하는 시점의 데이터를 통해 최적화와 학습이 이루어 지는것이 중요하다는 것을 확인할수 있었다.

5장에서 탄중변화에 따른 예측 성능을 확인하기 위해 장기 데이터를 취득하여 각 모델을 최적화 하고 예측 성능을 검증하는 실험을 수행하였다. 탄중 변화에도 NARX와 LSTM 최적화 모델의 예측 결과 값이 측정 값을 잘 반영하였으므로 석탄화력 발전소 공정운전에 적용 가능하다는 것을 확인할 수 있었다. 실험 결과는 60초와 120초 데이터 모두 LSTM 모델이 NARX 모델보다 더 우수한 예측 성능을 보여주었다.

본 논문에서 석탄화력 발전소 운전 데이터와 순환신경망을 활용하여 황산화물 배출 농도 예측 모델을 제안하였다. 예측 성능을 검증하는 실험을 통해 LSTM 최적화 모델이 예측 정확도가 높고 공정 운전에 적용될 수 있음을 확인할 수 있었다.

탈황공정의 특성에 맞추어 60초와 120초 샘플링 시간 간격을 갖는 데이터로 실험을 하였으나 다른 공정에 적용하기 위해서는 공정 특성에 맞는 변수와 샘플링 시간을 도출하는 연구가 필요하다. 예측 정확도를 높이기 위해서 입력 변수의 차원을 증가하거나 학습 데이터의 기간을 더 증가시키는 등 최적화 방안에 관한 연구도 수행할 수 있을 것이다.

본 논문은 순환신경망 모델의 예측 성능에 중점을 두고 연구를 수행하였으나 공정에 적용하기 위해서는 모델의 학습 시간과 데이터 예측에 걸리는 시간 확인 등 추가적인 연구가 필요하다. 이때 LSTM 모델과 예측 성능은 유사하나 더 간략화 된 모델인 GRU와의 비교 연구나 GPU를 연산에 활용하여 계산시간을 단축시키는 연구 등을 통해 석탄화력 발전소의 다양한 공정에 인공 신경망을 더욱 실질적으로 활용할 수 있을 것이다.

참고 문헌

- [1] G. Cheng and C. Zhang, “Desulfurization and denitrification technologies of coal-fired flue gas.,” *Polish Journal of Environmental Studies*, vol. 27, no. 2, 2018.
- [2] 환경부, “2019 제2차 기후변화대응 기본계획,” tech. rep., 관계부처 합동, 2019.
- [3] 환경부, “대기관리권역의 대기환경개선에 관한 특별법,” tech. rep., 환경부, 2020.
- [4] 환경부, “제2차 수도권 대기환경관리 기본계획 2015-2024 수정계획,” tech. rep., 환경부, 2020.
- [5] J. Lim, Y. Choi, G. Kim, H. Song, and J. Kim, “Modeling of wet flue gas desulfurization process for utilization of low-grade limestone,” *Korean Chemical Engineering Research*, vol. 57, no. 5, pp. 743–748, 2019.
- [6] J. Lim and J. Kim, “Optimization of a wet flue gas desulfurization system considering low-grade limestone and waste oyster shell,” *J. Korea Soc. Waste Manag*, vol. 37, pp. 1–12, 2020.
- [7] J. Fu, H. Xiao, T. Wang, R. Zhang, L. Wang, and X. Shi, “Prediction model of desulfurization efficiency of coal-fired power plants based on long short-term memory neural network,” in *2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pp. 40–45, IEEE, 2019.
- [8] J. Fu, H. Xiao, H. Wang, and J. Zhou, “Control strategy for denitrification efficiency of coal-fired power plant based on deep reinforcement learning,” *IEEE Access*, vol. 8, pp. 65127–65136, 2020.

- [9] H.-M. Jeong and J. H. Park, “Short-term electric load forecasting in winter and summer seasons using a narx neural network,” *The Transactions of the Korean Institute of Electrical Engineers*, vol. 66, no. 7, pp. 1001–1006, 2017.
- [10] B.-K. Jeon, K.-H. Lee, and E.-J. Kim, “Development of a prediction model of solar irradiances using lstm for use in building predictive control,” *Journal of the Korean Solar Energy Society*, vol. 39, no. 5, pp. 41–52, 2019.
- [11] M. J. Cheon and O. Lee, “A study on the stock price prediction and influence factors through narx neural network optimization,” *Journal of the Korea Academia-Industrial cooperation Society*, vol. 21, no. 8, pp. 572–578, 2020.
- [12] H. Xie, H. Tang, and Y.-H. Liao, “Time series prediction based on narx neural networks: An advanced approach,” in *2009 International conference on machine learning and cybernetics*, vol. 3, pp. 1275–1279, IEEE, 2009.
- [13] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, “Short-term residential load forecasting based on lstm recurrent neural network,” *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2017.
- [14] 김상연, *Time Series Data Analysis using Deep Learning in Industry*. PhD thesis, 서울대학교 대학원, 2019.
- [15] 이수지, *순차적 데이터의 회귀분석을 위한 순환신경망 훈련방법 연구*. PhD thesis, 서울대학교 대학원, 2016.
- [16] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [17] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

- [18] S.-N. Cheon, “화력발전소 미세먼지 배출현황과 저감,” *전기의세계*, vol. 67, no. 10, pp. 9–16, 2018.
- [19] Y. Guo, Z. Xu, C. Zheng, J. Shu, H. Dong, Y. Zhang, W. Weng, and X. Gao, “Modeling and optimization of wet flue gas desulfurization system based on a hybrid modeling method,” *Journal of the Air & Waste Management Association*, vol. 69, no. 5, pp. 565–575, 2019.
- [20] Y. Gong and Z.-G. Yang, “Corrosion evaluation of one wet desulfurization equipment–flue gas desulfurization unit,” *Fuel Processing Technology*, vol. 181, pp. 279–293, 2018.
- [21] P. Córdoba, “Status of flue gas desulphurisation (fgd) systems from coal-fired power plants: Overview of the physic-chemical control processes of wet limestone fgds,” *Fuel*, vol. 144, pp. 274–286, 2015.
- [22] 오일석, *기계학습(Machine Learning)*. 한빛아카데미, 2017.
- [23] 김성필, *딥러닝 첫걸음*. 한빛미디어, 2016.
- [24] S. Goki, *밑바닥부터 시작하는 딥러닝(Deep Learning from Scratch)*. 한빛미디어, 2017.
- [25] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O’Reilly Media, 2019.
- [26] S.-M. Lee, Y.-G. Sun, J. Lee, D. Lee, E.-I. Cho, D.-H. Park, Y.-B. Kim, I. Sim, and J.-Y. Kim, “Short-term power consumption forecasting based on iot power meter with lstm and gru deep learning,” *The Journal of The Institute of Internet, Broadcasting and Communication*, vol. 19, no. 5, pp. 79–85, 2019.
- [27] P. Tan, B. He, C. Zhang, D. Rao, S. Li, Q. Fang, and G. Chen, “Dynamic modeling of nox emission in a 660 mw coal-fired boiler with long short-term memory,” *Energy*, vol. 176, pp. 429–436, 2019.
- [28] 류근관, *통계학*. 범문사, 2010.

Abstract

Prediction Model of Desulfurization Efficiency of Coal-Fired Power Plants Based on RNN

Mi Yeon Seo

Graduate School of Practical Engineering

Seoul National University

Predicting sulfur dioxide(SO_x) emissions from flue gases in coal-fired power plants is challenging because flue gas desulfurization(FGD) performance is dependent on a wide range of variables. Prediction of SO_x emission amount is useful because it can be applied to process control and alarm circuits. The aim of this study was to forecast SO_x emissions from the FGD process in coal-fired power plants.

A novel prediction model based on an Recurrent Neural Network(RNN) herein was developed to forecast SO_x emissions. Operational data from coal-fired power plants were collected and divided into two sets for model training and testing. The author trained the proposed model for optimization using this training data set and verified its performance with test data set.

The SO_x emission prediction experiments showed that a suitable model must be selected and optimized using real operational data. The result of experiments showed that the LSTM optimization model was better at predicting emissions than the NARX model. The prediction of desulfurization efficiency based on RNN can provide important guidance for the operation and control of the FGD system.

In this study, the proposed prediction model was optimized using the operational data of an existing coal-fired power plant. The result of coal change experiments confirmed that the LSTM optimization model was suitable for application to coal-fired power plant process operations.

Keywords : Time-Series Data Prediction, FGD, NARX, LSTM, RNN

Student Number : 2019-24319