Tampere University

Sofia Koivumäki

DATA-DRIVEN MODELING OF CIRCULATING FLUIDIZED BED BOILER AIR EMISSIONS

Faculty of Engineering and Natural Sciences

Master of Science Thesis

May 2020

ABSTRACT

Sofia Koivumäki: Data-driven modeling of circulating fluidized bed boiler air emissions
Master of Science Thesis
Tampere University
Energy and Biorefining Engineering
May 2020

The tightened emission regulations and a continuous change towards low-grade fuel fractions make feasible air emission reduction more challenging in power plants. Therefore there is a need for data-driven advisor applications that can extract important information from process data in real-time. The objective of this thesis was to provide data-driven models of CFB boiler air emissions suitable for the use of the proposed cloud-based advisor system.

Based on the literature review models were implemented with multilayer perceptron method for SO_2, NO_x, CO emissions and costs. Models were trained with continuous process data gathered from a large scale reference multi-fuel CFB plant during its real operation. Process data from air, fuel and additives feed, bed and chamber conditions, process steam and flue gas stack measurements were used in modeling.

Model parameters were selected in four-fold cross-validation, in which the performance not only evaluated based on the prediction accuracy but also on the consistency between the modeled correlations and ones presented in the literature. Results show that models can predict emissions with satisfying accuracy and provide correlations between emissions and operational variables consistent with the literature in studied operation points.

It is likely, that significant improvements in the prediction accuracy cannot be achieved with this method without improving data quality and coverage, especially for fuel quality and biomass mixture. It was found out that the prediction accuracy of the models seemed to be more dependent on the process conditions than the model structure. Therefore it is suggested that models should be retrained often enough to maintain the prediction accuracy in everyday use.

To conclude this study is a solid base point for emission advisor cloud application models development. However, guarantee accurate and reliable model performance in varying process conditions, those should be developed further.

Keywords: CFB, emissions, modeling, MLP, neural networks

TIIVISTELMÄ

Sofia Koivumäki: Kiertoleijukattilan päästöjen dataan perustuva mallinnus
Diplomityö
Tampereen yliopisto
Energia- ja biojalostustekniikka
Toukokuu 2020

Kustannustehokas päästöjen hallinta on yhä haastavampaa voimalaitoksilla johtuen tiukentuneista päästörajoituksista sekä entistä huonompien polttoainejakeiden polttamisesta. Tämä luo tarpeen jatkuvatoimisille päästöjenhallintasovelluksille, jotka osaavat reaaliaikaisesti erottaa olennaisen tiedon suuresta määrästä prosessidataa. Tämän opinnäytetyön takoituksena on toteuttaa tietoohjautuvat mallit kiertoleijukattilan ilmanpäästöistä, joita voitaisiin hyödyntää osana pilvipohjaista päästöjenhallintasovellusta.

SO_2, NO_x, CO päästöjen sekä käyttökustannuksien mallinnuksessa käytettiin monikerroksinen perseptroniverkko -menetelmää, joka valittiin kirjallisuuskatsauksen perusteella. Mallit opetettiin jatkuvatoimisella prosessidatalla, joka oli kerätty suuren kokoluokan monipolttoaine kiertoleijukattilalaitokselta sen normalin toiminnan aikana. Malllinnuksessa käytettiin prosessidataa ilman- ja polttoianeen syötöstä, lisäaineiden syötöstä, palamisolosuhteista, prosessihöyrystä ja savuaaksun savupiippumittauksista.

Mallin parameterit valittiin neljän joukon ristiinvalidoinnilla. Ristiinvalidoinnissa parametrien vaikutusta sekä mallin ennustustarkkuuteen että maallinnettujen ja kirjallisuudessa esitettyjen korrelaatioiden yhdenmukaisuuteen. Tulokset osoittavat että mallit ennustivat päästöjä tyydyttävällä tarkkuudella. Tutkituissa prosessin ajotilanteissa mallit tuottivat korreaatioita päästöjen ja prosessin käyttömuuttujien välille, jotka olivat hyvin yhteneviä kirjallisuudessa esitettyjen korrelaatioiden kanssa.

On todennäköistä, että merkittäviä parannuksia mallien ennustamistarkkuudessa ei voida kyseisellä metodilla saavuttaa, jos ilman opetusdatan laadun parantamista ja laajuuden lisäämistä . Erityisesti polttoaineen laadusta ja biomassan koostumuksesta tulisi olla jatkuvatoimista dataa. Työssä huomattiin, että mallien suorituskyky vaikutti riippuvan enemmän prosessiolosuhteista kuin mallin parametreistä. Tämän perusteella suositellaan, että jokapäiväisessä käytössä oleva malli tulisi opettaa uudelleen riittävän usein.

Lopuksi voidaan todeta, että tämä opinnäyte työ luo hyvän pohjan päästöjenhallintasovelluksen mallien kehittämiseksi. Jotta mallien tarkkuus ja luotettavuus vaihtelevissa prosessiolosuhteissa voitaisiin taata, tulisi niitä edelleen kehittää.

Avainsanat: kiertopetikattila, päästöt, mallinnus, MLP, neuroverkot

PREFACE

This Master's Thesis was done for the Energy R&D department in Valmet Technologies Oy. I want to equally thank the whole energy R&D department and all colleagues who have shared their expertise and supported me during the process. Special thanks to Marko Palonen for giving me the opportunity to write my thesis in the familiar team, Kalle Vikkula and Tuukka Harmaala, for the support with data and modeling and Tuomas Petänen for giving the customer-oriented view to my thesis.

I want to thank Tero Joronen for suggesting this interesting subject for my thesis and being my company-side supervisor. I am grateful for all the expertise, support and great ideas he shared with me during the process. I am sincerely thankful for University lecturer Henrik Tolvanen at Tampere University. His guidance helped me to outline the big picture and overcome difficulties during the whole process.

I want to thank all my friends for the unforgettable memories and address a remarkable thanks to my brilliant friend and colleague Saara Väänänen for your great empathy and adventures. I would like to express my gratitude to my parents, sister, and especially Jaakko, for the all selfless support during my studies.

Tampereella, 16th May 2020

Sofia Koivumäki

CONTENTS

1 INTRODUCTION 1
2 CIRCULATING FLUIDIZED BED BOILER EMISSIONS 4
2.1 Circulating Fluidized Bed Boiler Components 4
2.2 Fuel Characteristics 7
2.3 NO_x Emissions 9
2.3.1 Formation 9
2.3.2 Reduction 11
2.4 SO_x Emissions 13
2.5 Emissions Cross Effects 15
2.6 Economics of Emission Control 18
2.7 Air Emission Regulation 19
3 DATA-DRIVEN MODELING OF BOILER EMISSIONS 21
3.1 Multilayer Perceptron 23
3.1.1 Structure 24
3.1.2 Training 27
3.1.3 Evaluation Metrics 31
4 MATERIALS AND METHODS 32
4.1 Data Selecting and Reprocessing 34
4.2 Model Input and Output Variables 36
4.3 Operational Variables 39
4.4 Neural Networks Training and Validation 41
5 RESULTS AND DISCUSSION 44
5.1 Emission Models Performance 44
5.1.1 SO_2 Model 45
5.1.2 NO_x Model 52
5.1.3 CO Model 58
5.1.4 Final Model Parameters 61
5.2 Overall Discussion 62
6 CONCLUSIONS 64
References 67

LIST OF FIGURES

2.1 CFB boiler hot loop (Valmet). 5
2.2 Flue gas system and auxiliaries for emission reduction – 1. flue gas duct 2. back-house filter 3. stack 4. induced draft fan 5. recirculation gas 6. limestone feeding 7. ammonia feeding. Adapted from (Valmet). 6
2.3 Combustion air system – 1. secondary air fan 2. primary air fan 3. primary air nozzles 4. secondary air nozzles. Adapted from (Valmet). 7
2.4 Nitrogen emission formation routes from fuel nitrogen in fluidized combustion. Combined from (Nussbaumer 2003; Raiko et al. 2002) 10
2.5 NO emission when fir/coal mix is burned in 8 MW CFB boiler. Adapted from (Leckner and Karlsson 1993) 11
2.6 NO emissions increase with bed temperature as well as with excess of air. Adapted from (Basu 2006) 12
2.7 Sulfur removal efficiency dependency on bed temperature in CFB boiler. Adapted from (Raiko et al. 2002) 14
2.8 The effect of varying air excess and temperature on NO_x, SO_2 and CO emissions and combustion efficiency adapted from (Lyngfelt, Åmand et al. 1998) 16
2.9 Cross-correlations cause challenges to emission control. (Leckner 1998) . 17
2.10 The effect of limestone addition to NO emission in 0.8 MW coal fired CFB boiler, Ca/S = 1.5. Adapted from (M. Hiltunen et al. 1991) 18

3.1 A model of a neuron. Adapted from (Haykin 2008) 24
3.2 Activation functions. Data gathered from (Glorot et al. 2011; Hastie et al. 2008) 25
3.3 An architectural graph of a MLP with two hidden layers. Adapted from (Haykin 2008). 26

4.1 Generalised schematics of the proposed emission advisor concept 33
4.2 Data was split into the training and validation data sets. 35

4.3 Selecting the final parameters for the emission models was an iterative process. 41

5.1 The effect of the number of neurons on validation scores in four-fold cross validation. 46
5.2 Trends of measured and predicted SO_2 emissions. Predictions are provided with the models 8-3-1 and 8-7-1. 47
5.3 Modeled correlations between SO_2 emission and operational variables. Slope of each line shows modeled correlation around an operation point (base point). 50
5.4 Generalized correlations presented in the literature. 50
5.5 The effect of the number of neurons on validation scores in four-fold cross validation. 53
5.6 Trends of measured and predicted NO_x emissions. Predictions are provided with the models 12-3-1 and 12-9-1. 54
5.7 Modeled correlations between NO_x emission and operational variables. Slope of each line shows modeled correlation around an operation point (base point). 57
5.8 Generalized NOx correlations presented in the literature. 57
5.9 The effect of the number of neurons on validation scores in four-fold cross validation. 59
5.10 A trend of measured and predicted CO emissions. Predictions are provided with the models 8-1-1 and 8-10-1. 60
5.11 Modeled correlations between CO emission and two operational variables. Slope of each line shows modeled correlation around an operation point (base point). 61
5.12 Generalized CO correlations presented in literature. 61

LIST OF TABLES

2.1 Typical properties of solid fuels. (Alakangas 2000; Houshfar et al. 2012; Raiko et al. 2002) 8
2.2 NO_x emission limit values (mg/Nm^3) for large combustion plants, calculated in 6 % O_2. (IED 2010) 20
2.3 SO_2 emission limit values (mg/Nm^3) for large combustion plants, calculated in 6 % O_2. (IED 2010) 20

4.1 Training and validation data sets 36
4.2 Variables used in the models. 37
4.3 Constants used in the cost calculation 39
4.4 Operational variables 40

5.1 The final model parameters 62

LIST OF ABBREVIATIONS

CO	Carbon Monoxide
HLN	Hidden Layer Neurons
N_2O	Nitrous Oxide
NH_3	Ammonia
NO_2	Nitrogen Dioxide
NO_x	Nitrogen oxides
NO	Nitric Oxide
SO_2	Sulfur Dioxide
SO_3	Sulfur Trioxide
CFB	Circulating Fluidized Bed
CHP	Combined Heat and Power
IED	The Industrial Emission Directive
MAE	Mean Absolute Error
MLP	Multi Layer Perceptron
PCA	Principal Component Analysis
PLS	Partial Lest Squares
R^2	Coefficient of Determination
RNN	Recurrent Neural Network
SCR	Selective Catalytic Reduction
SNCR	Selective Non-Catalytic Reduction

LIST OF SYMBOLS

Sign	Description	Unit
a_i	A neural network output	-
b_i	A neural network bias	-
C_{add}	Total cost of additives	€/kg
C_{aux}	Total cost of auxiliary power consumption	€/MWh
$\hat{C}$	An approximation of cost function	-
C_{loss}	Total cost of boiler heat loss	€/MWh
C_o	Operational cost	€/MWh
C	Cost function	-
δ	Sensitivity of the cost function	-
η	A learning rate	-
L	A neural network layer	-
m_{add}	An additive mass flow	kg/h
μ	A mean of training samples	-
P_{fuel}	Boiler fuel load	MW
p_f	Fan power consumption	€/MWh
p_{add}	An unit price of an additive	€/kg
p_e	Electricity price	€/MWh
p_f	Fuel price	€/MWh
σ	Standard deviation of training samples	-
$\sigma()$	An activation function	-
v_i	A linear combination of inputs and weights	-
w_{ij}	A neural network weight	-

Sign	Description	Unit
x_j	A neural network input	-
$\mathbf{x_k}$	Neural network inputs	-
$\bar{y}$	A mean of observed outputs	-
$\hat{y}_i$	A Predicted output	-
y_i	An observed output	-
$\mathbf{y_k}$	Desired outputs	-

1 INTRODUCTION

Combustion processes in energy and industrial sectors are one of the primary sources of harmful air pollutants, such as nitrogen oxides, sulfur dioxide, carbon monoxide and particulate matters. These pollutants are harmful to both environment and humans since they cause acid rain, photogenic smog and breathing issues, for example. In recent years, the interest in protecting the environment from these harmful pollutants has been increasing (Ministry of the Environment 2019). As a result, authorities have tightened the emission regulations. For example, in 2016, the European Union has set the new industrial emissions directive (IED), which tightened the flue gas emission limits for NO_x, SO_2 and dust emissions in the industrial combustion plants (IED 2010).

The climate change mitigation has led to transformations in energy production. For example, several reports have presented biomass combustion to play an increasing role in the low-carbon energy scenarios and the continuous changeover towards more complex and challenging fuel fractions, such as agricultural residues and industrial waste, is likely to occur (Alakangas et al. 2018; IEA 2017). Thus, the utilization of varying fuel mix comes with consequences, as noted by Hupa (2005), The diversity of the fuel types makes emission control challenging, since the air emissions are strongly dependent on the fuel. Besides that, a growing share of weather-dependent renewable energy sources has changed the behavior of electricity markets (Finnish Energy 2019). Not only the average price of electricity has declined, but also the short-term variance has increased in the stock prices (Fingrid 2017).

As a result, the operation of conventional combined heat and power (CHP) plants has become more and more challenging, since those need to meet the strict emission targets, tolerate widely varying fuel quality and respond fast to fluctuating load demands and electricity prices, simultaneously. Also, all these should be carried out with feasible operational costs and high efficiency.

At the same time, cloud computing has transformed the way process and energy industries deploy process data in their daily operations. Cloud computing allows cost-efficient storage of large data vaults and provides scalable computing resources for data processing and analysis. It also enables easy access to data from anywhere, not only from the local on-site connection. This, together with the deployment of advanced data-based modeling algorithms, provides an opportunity to find out the hidden correlations and dependencies in data and monitor those in real-time (Martinsuo et al. 2018). Therefore solutions that deploy data-driven modeling of power plant emissions may help the energy suppliers to meet the tightened emission targets in the challenging operating environment in a feasible way.

In the literature, several different approaches have been proposed for data-driven modeling of power plant emissions, however only a few for biomass combustion. Both M et al. and Krzywanski et al. have focused on modeling circulating fluidized bed (CFB) combustion emissions: M et al. have studied an online decision-supporting system for the NO_x emission efficiency improvement in biomass and coal co-combustion, (2011a), (2011b) and (2016), whereas Krzywanski et al. has presented a generalized models for NO_x (2017) and SO_2 prediction in coal combustion (2014). Further, Korpela et al. (2017) have presented satisfying results for indirect monitoring of NO_x emissions in a natural gas boiler with data-driven methods and Golgiyaz et al. (2019) has implemented flame image-based models for multiple emissions in coal firing boiler. Although extensive research has been carried out on data-driven emission modeling, not many studies exist, which takes account the multiple emissions at once.

This study aims to examine, which data-driven method would be suitable for CFB combustion emission modeling, and then model NO_x, SO_2 and CO emissions selected method. Models are developed to be suitable for a cloud application framework, which advises operators on how to change certain operational variables to achieve acceptable emission levels in a feasible way, in real-time. The implementation of multi-variate optimization and a cloud application is beyond the scope of this work. However, those are taken into account in model development. Therefore only continuous process data form the reference plant available in Valmet cloud have been used for modeling.

The work offers an important insight into the capabilities and limitations of data-driven emission modeling with the continuous process data gathered from the multi-fuel CFB boiler.

Following research questions will be discussed:

1. What is a suitable data-driven method for modeling CFB boiler emissions?
2. Which continuous process variables are needed for modeling?
3. Which model parameters provide the best prediction accuracy in examined case?
4. Are the correlations between selected operating variables and modeled emissions of this case consistent to the ones presented in the literature?
5. Are proposed emission models accurate enough to be used in cloud application?

The first two research questions are discussed on a theoretical basis in Chapters 2 and 3. Chapter 2 presents an overview of CFB combustion. The main focus is on the air emission formation and reduction in biomass or co-combustion. Chapter 3 provides a brief summary of data-driven methods presented in the literature for power plant air emissions modeling. Further, it offers a more comprehensive approach to multilayer perceptron (MLP) regression with backpropagation training.

Chapter 4 describes the data selection from the reference plant and its preprocessing methods and presents MLP neural network training and validation process for three emission models. Further, the operational parameters and operational cost calculations for the cost model are revealed. The results for prediction accuracy with different model configurations and the correlation between operational variables and modeled emission are presented and discussed individually for each emission in chapter 5. Finally, chapter 6 outlines the most important findings and addresses the final research question. It also proposes further work in the field of the study.

2 CIRCULATING FLUIDIZED BED BOILER EMISSIONS

CFB combustion is a widely employed technology for the large scale power generation from solid fuels. CFB technology belongs to the family of fluidized bed technology and it states back to 1970 when the first commercial CFB boiler started to operate. Since then, the installed capacity of CFB boilers has increased steadily, according to Transparency Market Research (2014) the installed capacity of CFB boilers was 92 GW in 2014 and Global Market Insights (2019) has estimated it to exceed 400 GW by 2025.

The combustion process in energy generation and industrial processes is one of the major sources of NO_x and SO_2 emissions, which are harmful to both environment and humans (Ministry of the Environment 2019). The development of CFB technology tends to offer an environmental friendly alternative to the grate combustion since it is known for its relatively low nitrogen oxides emission levels and a possibility for cost-efficient in-furnace sulfur capture. These advantages are base on the uniform combustion temperature, approximately 850 °C, and the efficient mixing of fuel and air during the combustion process, which circulating fluid ensures. The other remarkable benefit of CFB technology is the ability to exploit the low-grade fuels such as forest residues, bark and industrial sludge. Also, it is capable of efficient multi-fuel combustion. This is a consequence of the large heat capacity of the bed and the circulation, ensuring efficient fuel mixing. (Raiko et al. 2002)

2.1 Circulating Fluidized Bed Boiler Components

In the circulating fluidized the combustion of solid fuel occurs in the suspension called fluidized bed, which circulates in the hot loop. The bed material is typically a mixture of granular solids, such as sand, fuel ash and a sulfur capture sorbent. The fluidization of the bed material is caused by air injection through the evenly distributed nozzles at the

bottom of the furnace. In CFB boiler, the fluidizing velocity needs to be high enough, up to 8 m/s, to ensure the circulation of the bed material circulate in the hot loop. (Basu 2006)

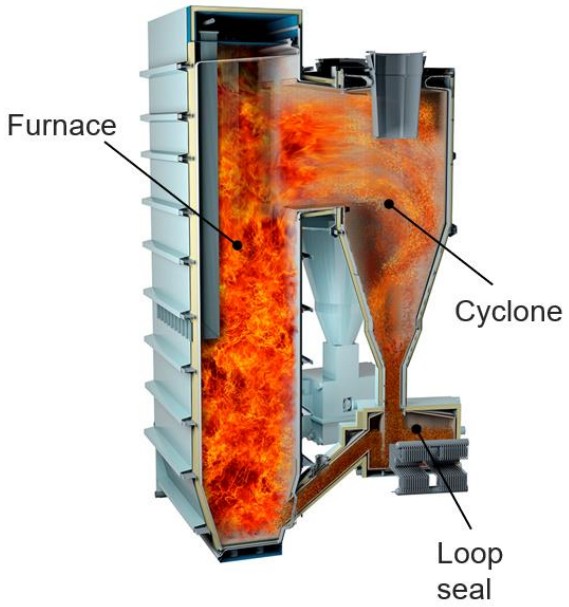


Figure 2.1. *CFB boiler hot loop (Valmet).*

Regardless of the manufacturer, the hot loop usually shares a basic configuration comprising: furnace, cyclone and loop-seal. These components are pointed out in figure 2.1, which illustrates the Valmet CFB boiler. In the recirculation process, the bed material propels from the lower part of the furnace to the upper part of it along the flue gas. Then the flue gas–bed material suspension enters the cyclone, where the circulating solid matter is separated from the flue gas. The solid matter is circulated back into the lower part of the furnace through the loop-seal and the flue gas flows forward in the back-pass, releasing the heat to the water-steam system. (Spliethoff 2010)

In addition to the combustion system, there are other important systems supporting the combustion process, such as water-steam, fuel feeding, air, flue gas and auxiliary, in the boiler scope (Rayaprolu 2009). In terms of the emission control, the most remarkable ones are the combustion air, flue gas handling and auxiliary systems, which are presented in the figures 2.3 and 2.2.

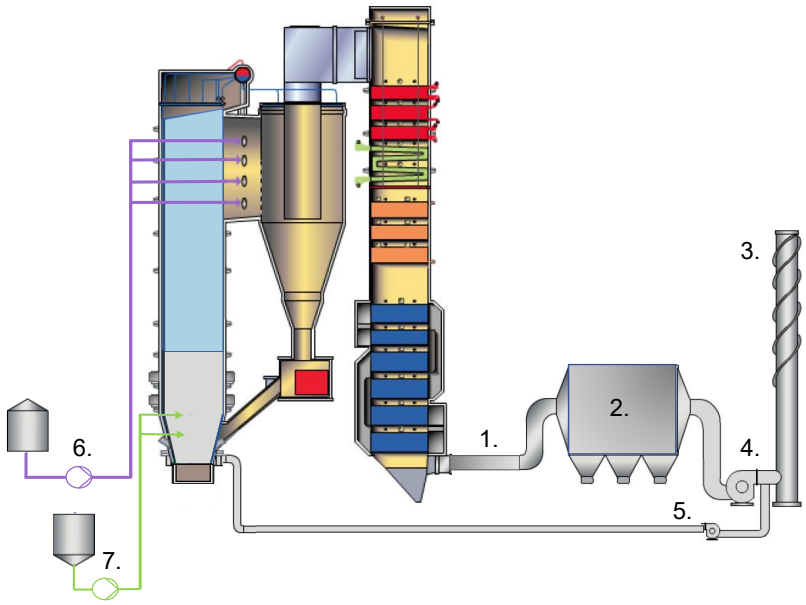


Figure 2.2. *Flue gas system and auxiliaries for emission reduction – 1. flue gas duct 2. back-house filter 3. stack 4. induced draft fan 5. recirculation gas 6. limestone feeding 7. ammonia feeding. Adapted from (Valmet).*

When flue gas is traveled trough all the heat surfaces, is it conducted to the flue gas handling system. Before that, a portion of flue gas may be recirculated back to the furnace. The main purpose of the flue gas handling system is to remove flue gases from the boiler, separate the solid matter of it and remove remaining NO_x and SO_2 emissions of it, if necessary. Figure 2.2 shows typical components of the flue gas handling system and also auxiliary equipment that are typically applied for emission reduction: storage and injection of ammonia and hydrated limestone. In general auxiliary systems comprises a group of processes and equipment, which support the boiler operation. (Koskelainen et al. 2007)

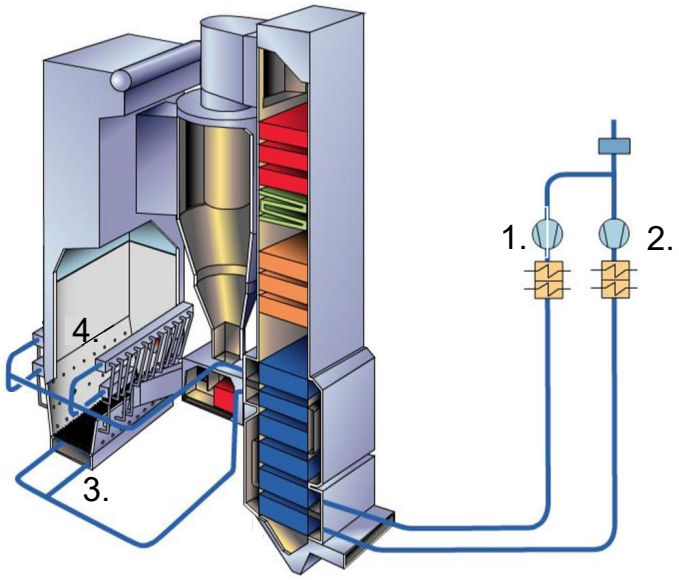


Figure 2.3. *Combustion air system – 1. secondary air fan 2. primary air fan 3. primary air nozzles 4. secondary air nozzles. Adapted from (Valmet).*

The combustion air system is presented in figure 2.3, which shoes that it consists of fans, ducts, nozzles and air preheaters. Its purpose is to provide an amount of the combustion air for efficient combustion and to fluidize the bed material. (Koskelainen et al. 2007) The combustion air is typically divided into primary and secondary airs to enhance the control of the combustion process and NO_x emissions generation. Primary air is fed to the boiler through the evenly distributed nozzles at the bottom of the furnace. It is the provider of the bed fluidizing and the main oxidizer for the combustion. Secondary air is fed to the furnace from the several points on the walls at the height lower tapered section of the bed. The purpose of secondary air is to finalize the combustion of solid fuel and its share is typically between 40–60 %. (Basu 2006)

2.2 Fuel Characteristics

Biomass is a typical fuel in the CFB boilers and its properties have a significant impact on the emission generation during the combustion process. Biomass covers a wide range of feedstocks, such as wood, bark, forest residues, sludges and straws (Raiko et al. 2002). The table 2.1 presents both chemical and physical properties to the typical biomass types. The same properties are also listed for coal to give a benchmark for biomass properties since most combustion technologies are originally designed for fossil fuel combustion.

Table 2.1. *Typical properties of solid fuels. (Alakangas 2000; Houshfar et al. 2012; Raiko et al. 2002)*

Property	Wood	Bark	Forest residues	Straw	Sludge	Coal
Moisture, %	30–45	40–65	50–60	2–12	60–80	10
Volatile Matters, % (d)	84–88	70–80	77–80	66–78	12–60	29
Ash, % (d)	0.4–0.5	1–3	2.3	5–6	45–97	14
C, % (d)	48–50	51–66	51–53	39–48	25–66	76–87
H, % (d)	6.0–6.5	5.9–8.4	6.0–6.2	5.0–6.0	4.0–7.0	3.5–5.0
O, % (d)	38–42	24–40	40–41	35–41	22–50	3–11
N, % (d)	0.5–2.3	0.3–0.8	0.4	0.8–1.7	<2.0	0.8-1.2
S, % (d)	0.05	0.05	0.02	<0.2	<1.5	<0.5

Therefore coal behavior is well known in the literature. The primary purpose of the table is to visualize the large varying of fuel properties, not only between the feedstocks but even within one since it tends to be a key challenge with the robust boiler operation and also often with the biomass modeling. polttojapalaminen

Hupa (2005) has studied the biomass conversion process in the combustion and compared it to the one of coal. Stages of the thermal conversion of combustion are drying, the release of volatiles and char conversion. Biomass has a relatively high share of volatile matters and a low share of fixed char (see table 2.1), which affects the conversion dynamics and further, to the emission generation routes. Those, together with the shares of elementary nitrogen and sulfur in the fuel, are the significant factories defining the emission generation of NO_x and SO_2 emissions in the combustion process.

It is typical that in the biomass combustion both, NO_x and SO_2 emissions are lower than in the coal combustion. However, the throwback in biomass combustion often seems to be the challenge with emission control and modeling. Even though biomass behavior and emission generation during the combustion process have been studied widely, there are still some unsolved issues. Because of its complex nature and greatly varying properties, the formation routes of NO_x emissions, especially during the co-combustion processes, still lacks a deep understanding. (Hupa et al. 2017)

2.3 NO_x Emissions

Combustion is one of the key sources of nitrogen oxides emissions, which are often considered the main air pollutants. NO_x emissions are responsible for several environmental issues, such as acid rain, photochemical smog, ozone layer depletion and troposphere ozone. The exposure to high concentrations of nitrogen oxides is also detected to cause many health issues. To decrease these detriments of NO_x emissions and reach the tightened emission limits in most efficient way NO_x formation and reduction have been studied a lot lately, especially in biomass combustion. The formation of NO_x is reduced by optimizing the combustion conditions, along with that several NO_x reduction technologies are widely used, such as selective non-catalytic reduction (SNCR) and selective catalytic reduction (SCR). (Skalska 2010)

A diverse selection of nitrogen oxides exists in the environment (Skalska 2010). The abbreviation NO_x generally refers to nitric oxide NO and nitroden dioxide NO_2, though. These are the main nitrogen oxides emitted form the combustion process, alongside nitrous oxide (N_2O). NO_x emissions in the flue gas contains approximately 95 % NO and only 5% NO_2. (Gomez-Garcia et al. 2015; Wang et al. 2007) However, the environmental impact of the both components in NO_x can be considered similar since most of the nitric oxide oxides to NO_2 in the atmosphere. (Raiko et al. 2002)

2.3.1 Formation

In the combustion process, nitrogen oxides form via oxidation of atmospheric nitrogen and fuel-bound organic nitrogen. The atmospheric nitrogen mainly reacts trough the thermal NO_x mechanism, which requires sufficiently high temperature, at least 1300 °C, to occur. When considering the CFB boiler, the combustion temperature is far too low for thermal NO_x to form. Therefore in the CFB combustion, the main source of NO_x is the oxidation of the fuel-bound nitrogen. (Raiko et al. 2002)

The general mechanism for NO_x to form from solid fuel nitrogen is illustrated in 2.4: During fuel pyrolysis, the volatile nitrogen compounds together with volatile carbon are released. Volatile nitrogen is typically released as hydrogen cyanide (HCN) or ammonia (NH_3). Thus, some of the fuel nitrogen and carbon remains in the solid char and the amount of released volatiles differs among the fuel. Abelha et al. (2008) studied that in the

fluidized bed combustion of woody biomass over 80 % of the nitrogen is released with the volatiles. It is also proven that biomass have higher NH_3/HCN ratio during devolatilization than coal (Konttinen et al. 2013). Despite the amount of released volatiles, those react further to NO in the presence of oxygen. Nitrogen may react further to N_2O or elementary nitrogen (N_2). (Raiko et al. 2002)

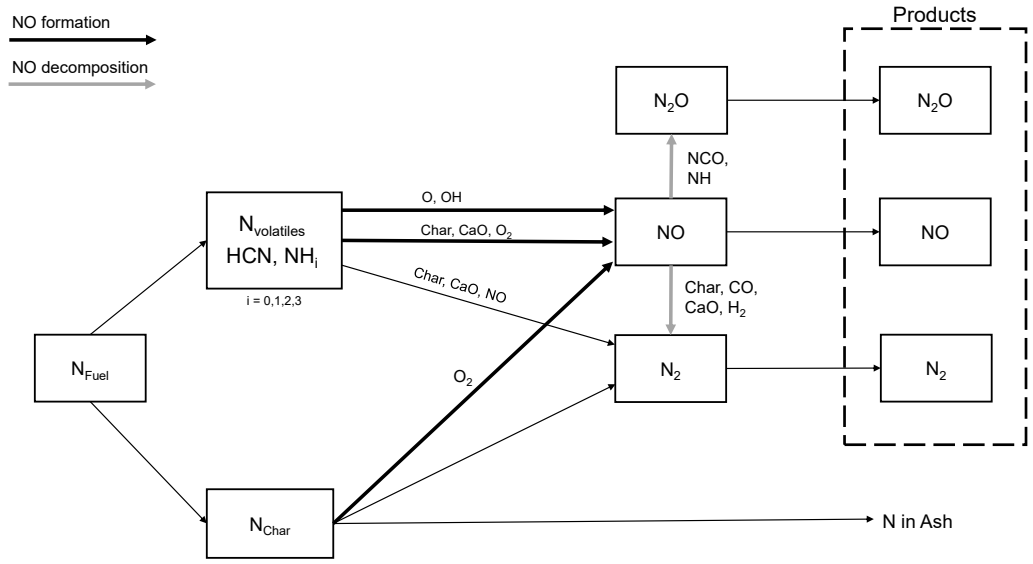


Figure 2.4. *Nitrogen emission formation routes from fuel nitrogen in fluidized combustion. Combined from (Nussbaumer 2003; Raiko et al. 2002)*

It is possible for formed NO to reduce back by char to elementary nitrogen, as seen in figure 2.4. Typically there is more char in the furnace during the coal combustion than biomass combustion (Raiko et al. 2002). This may lead to the situation presented by Leckner, Åmand et al. (2004): The NO_x emissions can be higher for wood-based fuels than coal, even though wood-based fuels contain significantly less nitrogen than coal. The reducing effect of char is also presented to cause the strong non-linearity between fuel type and nitrogen oxide emissions in co-combustion 2.5 (Leckner 2007). All in all, the dependency between emitted NO_x and fuel type is very complex and it is still under the research (Konttinen et al. 2013; Leckner, Åmand et al. 2004; Vermeulen et al. 2012).

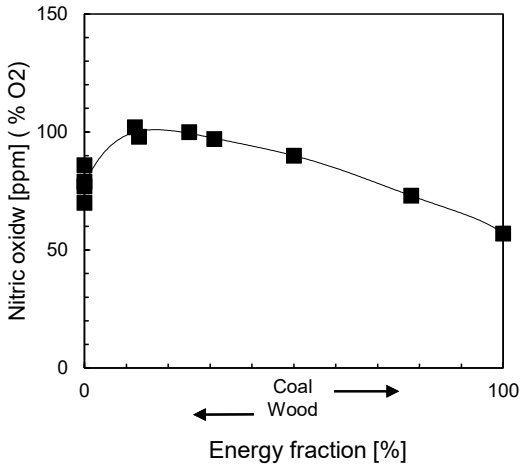


Figure 2.5. *NO emission when fir/coal mix is burned in 8 MW CFB boiler. Adapted from (Leckner and Karlsson 1993)*

2.3.2 Reduction

The optimization of combustion conditions is often presented to be a primary way to control the NO_x emissions in CFB combustion. The air excess ratio, together with the combustion temperature, has proven, in several studies and commercial applications, to have a major effect on NO formation. Figure 2.6 shows the correlation of nitrogen emissions to the combustion temperature and air excess. When the total excess air ratio grows, the oxidizing circumstances increase, which onward linearly increases the nitrogen oxide content. The change in the combustion temperature does not influence the relation of air excess and nitrogen oxides. However 2.6 reveals that the low combustion temperature indicates low NO_x emissions. (Basu 2006; Raiko et al. 2002)

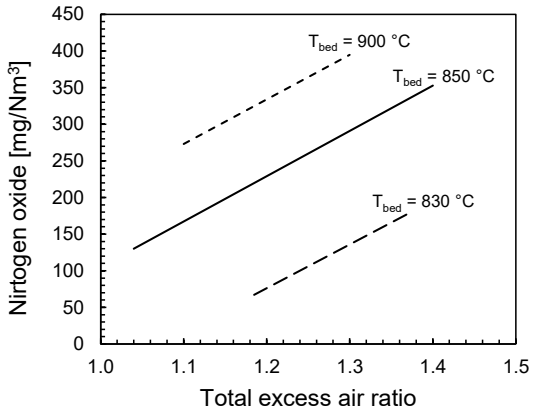


Figure 2.6. *NO emissions increase with bed temperature as well as with excess of air. Adapted from (Basu 2006)*

Typically combustion conditions, and thereby the NO_x emissions formation are controlled with the air staging in the fluidized bed boilers. This means that the combustion air is divided into primary and secondary airs and the feeding of them is staged. The general idea of the air staging is to keep the primary combustion region slightly air-deficient, which reduce the probability of the oxidation reactions of volatile matters (NH_3 and HCN) to the nitrogen oxide. For CFB combustion, a less significant benefit of air staging is that it also reduces the combustion temperature, which onward reduces the formation of thermal NO_x. Further, air staging does not only decrease the NO formation reactions, but it also forms more reducing zones to the lower regions of the furnace, which enables the NO reduction reaction back to the nitrogen in the presence of char or CO. The completeness of air staging can be evaluated with the primary air ratio, which expresses the ratio between primary air and secondary air. It is often challenging to reach a complete air staging at the low boiler loads since a certain amount of primary air is needed to ensure the bed fluidization. Therefore, the optimal primary air ratio may not be reached. (Basu 2006) Leckner (1998) and Qian et al. (2011) have both studied that flue gas recirculation often decreases nitrogen oxide emissions. Qian et al. (2011) presents that flue gas recirculation makes the residence time of nitrogen oxides longer in the furnace, which increases the chance of its further reduction. Also, flue gas recirculation reduces the O_2 content in the furnace, which leads to the lower NO generation.

At some process states, NO_x emissions may not be reduced enough only by affecting the combustion conditions. In that case, NO_x emissions can be reduced with secondary methods, such as selective catalytic reduction and selective non-catalytic reduction of

nitrogen oxides. The use of catalytic increases NO removal efficiency in SCR up to 90 %, but makes it also a more expensive solution when comparing to SNCR. The NO_x removal efficiency of SNCR is from 40 % to 70 % . (Valmet 2019) In SNCR, ammonia NH_3 is injected to the flue gas to reduce NO_x, and therefore combustion conditions affect it's performance, unlike in SCR. In the CFB boiler, ammonia is typically injected into the upper part of the furnace or to the cyclone inlet. (Basu 2006) In SNCR the NO_x reduction is mainly based on the following reaction

$$NH_3 \xrightarrow{+OH,+O} NHi \xrightarrow{+NO} N_2 \quad \{1\}$$

In which ammonia decompose and reduce in the presence of oxygen and hydroxide radicals. Reaction 1 is very sensitive to temperature and the optimal temperature for NO_x reduction is around 900 °C. At high temperatures, added ammonia may produce NO instead of reducing it. On the other hand, if the temperature gets too low, ammonia will not react fast enough and high ammonia slip is detected fro the flue gas. The performance efficiency of SNCR is not only dependent on temperature but also on CO content since CO burns in the cyclone, causing the production of radicals and high local temperatures, both influencing the SNCR reactions.

2.4 SO_x Emissions

SO_2 and SO_3 emissions are harmful to both environment and human health. For example, high sulfur oxides concentrations in the atmosphere cause acid rain. It leads to the acidification of the ecosystem and irritate human airways. (Ministry of the Environment 2019) In CFB combustion SO_2 emissions are mainly originated from the fuel sulfur. It is typical that all sulfur in the fuel forms either SO_2 or SO_3, however the SO_2 formation is large compared to SO_3 formation. Due to this, formed SO_2 varies a lot between different fuel types. For example, sulfur content, and therefore SO_2 emissions, can be even ten times higher for coal than for biomass. (Raiko et al. 2002)

During the combustion process, typically nearly all fuel bound sulfur releases and generates SO_2. Therefore desulfurization is often applied to reach emission the targets of emission regulation. In the CFB boiler, desulfurization can be done in situ by adding a solid sorbent to the furnace within the bed material (**Gungor**; Raiko et al. 2002).

The most typical and cost-efficient sorbents added are limestone (CaCO3) and dolomite ($CaCO_3 \bullet MgCO_3$), but Mathieu et al. (2013) has studied the potential of several other sorbents as well. The sulfur absorption in the furnace occurs via two main reactions: limestone calcination 2 and desulfurization 3

$$CaCO_3 \longrightarrow CaO + CO_2 \quad \{2\}$$

$$CaO + SO_2 + \frac{1}{2} O_2 \longrightarrow CaSO_4 \quad \{3\}$$

The bed temperature needs to exceed 800 /degree C for calcination reaction to occur whereas desulfurization reaction takes place within a larger temperature range. However, the increase in temperature may decrease the sulfur removal efficiency in CFB, since the equilibrium between CaO, $CaSO_4$ and CaS changes. In the desulfurization reaction SO_2 is absorbed to $CaSO_4$ if the combustion conditions are oxidizing and to CaS if the conditions are reducting. Further, if the particle emerges to reducing zones, calcium sulfate may reduce back to SO_2 (Hansen et al. 1993). This implies that sulfur removal is sensitive to temperature changes and according to several studies, the optimal temperature range for sulfur removal is from 800 to 850 ℃, as presented in figure 2.7 (Ahrlich 1975; Raiko et al. 2002; Tarelho et al. 2005).

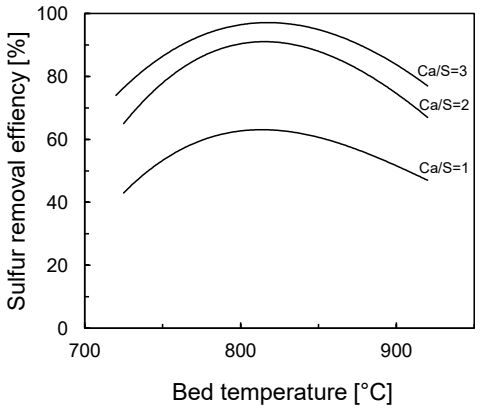


Figure 2.7. *Sulfur removal efficiency dependency on bed temperature in CFB boiler. Adapted from (Raiko et al. 2002)*

The figure also shows that the increase in Ca/S molar ratio enhances the capture efficiency and typically the optimal range for it is between two and three (Basu 2006). Lyn-

gfelt and Leckner (1999) and Anthony et al. (2001) have found that to take benefit from all the added limestone, the particle size should be small enough to ensure a large reactive surface. Also, a long residence time, for example, via flue gas recirculation, increases the sulfur reduction. Fuel characteristics do not only define the amount of formed SO_2 but also have an impact on sulfur removal efficiency. Lower sulfur content in fuel implies a lower sulfur capture rate since the SO_2 concentration in combustion is smaller and, therefore, more difficult for limestone to absorb. Besides the sulfur content, also the calcium content in the fuel ash matters since it determinate the rate of self-absorption. If the calcium content in the ash is high, the need for additional limestone feed is smaller. The share of the volatiles affect the combustion dynamics and thus may determinate the place where sulfur is released in combustion. (Raiko et al. 2002)

Also, both air excess ratio and air staging have detected to correlate with sulfur removal rate in several studies. Since the desulfurization reaction takes place in oxidizing conditions, low air excess ratio often reduces the sulfur removal efficiency. Thus it increases reduction zones. Mainly for the same reason, the strong air staging, which means a low share of the primary air, is studied to reduce the desulfurization rate. (Gungor 2008; Tarelho et al. 2005)

2.5 Emissions Cross Effects

In the sections 3.1.3 and 2.4 emission reduction methods are examined individually for the both emissions. However, it is a well-known fact that emission control methods tend to have advantageous effects on the one emission level disadvantageous ones on the other. When considering the overall performance of the CFB boiler, it is not only important to examine how does the reduction methods influence the other emissions, but also the overall combustion process and its efficiency. It is also typical, that reduction of emission generation during the combustion may cause additional CO emissions, mainly due to the limited excess of air (Lyngfelt, Åmand et al. 1998). This is good to take account, even though, according to Lyngfelt and Leckner (1999), CO emissions are not usually perceived to be a major problem under normal operating conditions in the CFB boiler. However, challenges with simultaneous control of NO and CO emissions are highlighted at the low boiler loads (Lyngfelt and Leckner 1999).

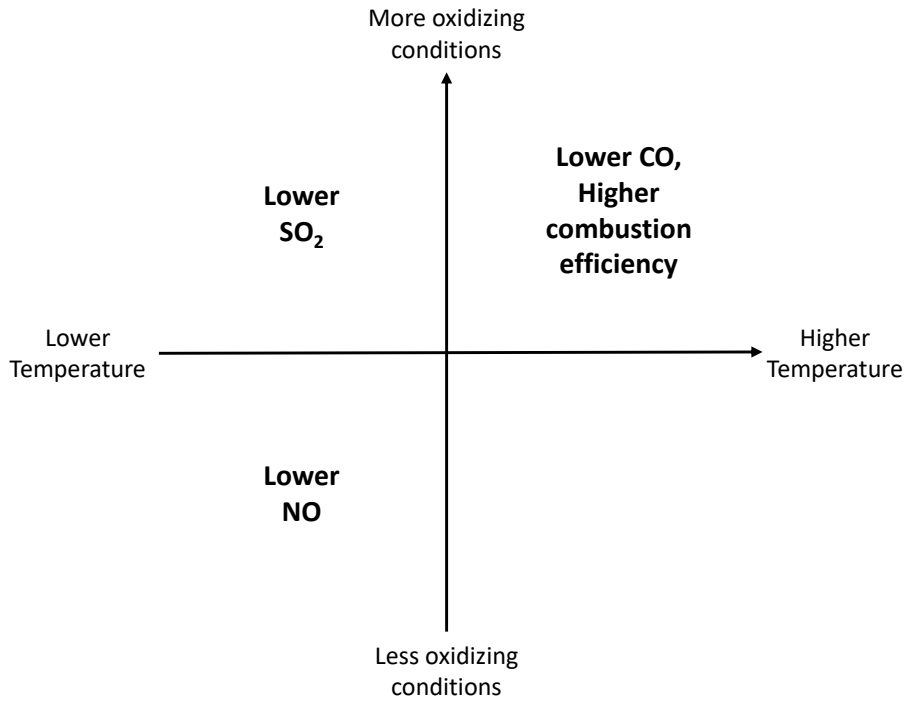


Figure 2.8. *The effect of varying air excess and temperature on NO_x, SO_2 and CO emissions and combustion efficiency adapted from (Lyngfelt, Åmand et al. 1998)*

Lyngfelt, Åmand et al. (1998) has studied the cross-correlation between the NO_x, SO_2, CO emissions behaviour in terms the state of the combustion temperature and air excess. He has also taken into account how does the state of those combustion parameters influence the combustion efficiency. These correlations are simplified to the coordinate system of temperature and air excess, which is in figure 2.8. The figure 2.8 emphasis the difficulty of the multi-emission control. It is possible to decrease nitrogen oxide emissions by lowering the combustion air excess. However, it is typical that SO_2 emissions and CO emissions levels may grow simultaneously. Further, the high CO content in the flue gas may reduce the efficiency of NO reaction to ammonia. Naturally, the low excess of air reduces the completeness of the combustion and therefore, under that condition, the combustion efficiency is smaller. When considering the combustion temperature, the reduction of SO_2 and NO_x emissions requires a lower temperature than would be optimal in terms of combustion efficiency and CO formation.

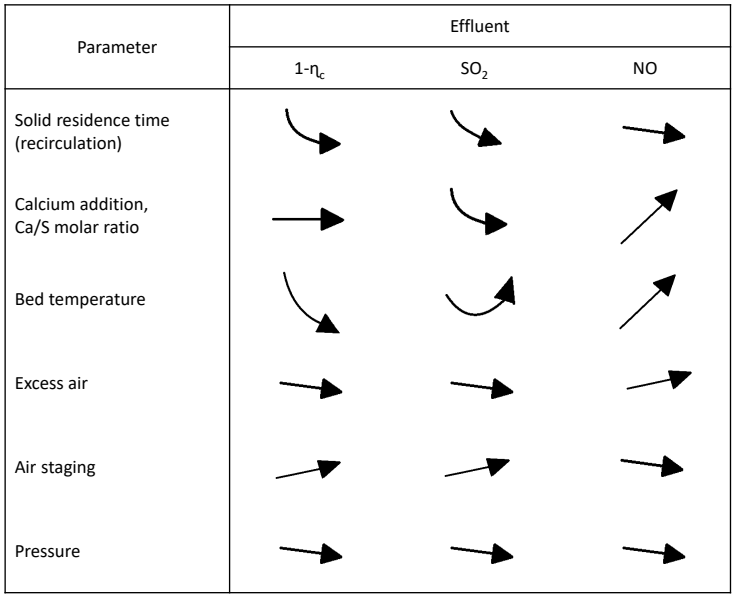


Figure 2.9. *Cross-correlations cause challenges to emission control. (Leckner 1998)*

Leckner (1998) has expressed the challenge of the multi-component emission control in CFB combustion. He has studied the simultaneous impact of six crucial emission control parameters to NO, SO_2 emissions and combustion efficiency. Those variables are reticulation, calcium addition, bed temperature, excess air, air staging and pressure. The figure2.9 presents, how does each effluent behave, if the control parameter grows - in general, if arrows point up, the impact is negative, if it points down the effect is positive. Leckner (1998) has reported results about the bed temperature and air excess that are in-line with Lyngfelt1998's observations. He presents that air staging is often applied for NO reduction with good results, however, as a drawback, it usually weakens the efficiency of sulfur capture and therefore increases SO_2 emissions.

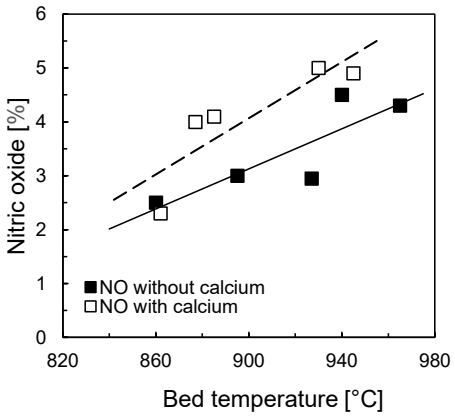


Figure 2.10. *The effect of limestone addition to NO emission in 0.8 MW coal fired CFB boiler, Ca/S = 1.5. Adapted from (M. Hiltunen et al. 1991)*

Also, the addition of limestone for sulfur capture have some unpleasant side-effects, since Leckner (1998) has noticed it to increase nitrogen oxides. M. Hiltunen et al. (1991) has studied this further in the 0.8 MW pilot CFB boiler, and gained converging results, which are presented in the figure 2.10.

2.6 Economics of Emission Control

In the emission control often the target is, not only reach the regulated emission levels, but also do it in as feasible way as possible. Emission reduction generates operational costs for CFB plant mainly via three routes: the cost of additives, auxiliary power consumption, losses in efficiency. Generally used additives in the emission control are limestone and ammonia, which have a quite substantial effect on total operating costs of CFB boiler. The cost of additives can be calculated as follow,

$$C_{add} = \sum_{i=1}^{N} \dot{m}_{add,i} p_{add,i} \tag{2.1}$$

where C_{add} [€/h] is the total cost of additives, $\dot{m}_{add,i}$ [kg/h] is mass flow of an additive and $p_{add,i}$ [€/kg] is a unit price of a additive. The unit price of additives depends on the market situation and location of the power plant.

The air system is one of the primary sources of auxiliary power consumption. Even though air emission control is not responsible for the whole auxiliary power consumption

of the fans, it often causes changes to the power consumption of primary air fans, secondary air fans, induced draft fan and recirculation fan which change the total auxiliary power consumption costs of CFB plant according to the equation

$$C_{aux} = p_e \sum_{i=1}^{N} P_{f,i} \tag{2.2}$$

where C_{aux} [€/h] is the cost of auxiliary power consumption of the fans, $P_{f,i}$ [MW] is the power consumption of a fan, such as primary air fan, secondary air fan, ID fan and recirculation fan, and p_e [€/MWh] is the price of electricity.

The last cost component in the operational costs of air emissions, heat loss, ties up the air emissions and boiler efficiency. Heat loss is calculated with

$$C_{loss} = (1 - \eta) \sum_{i=1}^{N} P_{fuel,i} p_{f,i} \tag{2.3}$$

where C_{loss} [€/h] is the cost caused by the heat loss in boiler, η is boiler efficiency, $P_{fuel,i}$ [MW] is a fuel load in the boiler and $p_{f,i}$ [€/MWh] is the price of corresponding fuel. The total operational cost is then

$$C_o = C_{add} + C_{aux} + C_{loss} \tag{2.4}$$

where C_o [€/h] is the operational cost.

2.7 Air Emission Regulation

In recent years the interest to protect the environment from harmful emissions has grown. As a result, authorities have introduced new environmental laws and emission regulation has tightened. European Union has set several directives to control air emissions. The latest is the directive on the reduction of national emissions of certain atmospheric pollutants (2016/2284/EU), which sets annual limits for nitrogen oxides, sulfur dioxide, ammonium, fine particulate matter and nonmethane volatile organic compounds in the national level. This directive obligates each member country to write a national air control program to reach the annual emission limits. (Ministry of the Environment 2019)

In addition to these national annual limits, the EU has set regulations concerning sectors

with high air emissions. It is well known that the combustion processes performing in industry and energy production sectors are still the major sources for NO_x, SO_2, fine particulate matter and volatile organic compounds emissions. These cause both health issues and environmental problems. (Ministry of the Environment 2019) The main EU instrument concerning the air emission control in energy production is the Industrial Emissions Directive (IED), which lay down on rules for integrated emissions prevention and control in the combustion process. IED sets tightened limits for flue gas emissions including NO_x and SO_2 emissions. Those limits are presented in the tables 2.2 and 2.3. IED is applicable in power plants, which exceed 50 MWth and have guaranteed the operational permit later than 7 January 2013. (IED 2010; Ministry of the Environment 2019)

Table 2.2. *NO_x emission limit values (mg/Nm^3) for large combustion plants, calculated in 6 % O_2. (IED 2010)*

Total rated thermal input (MW)	Coal and lignite other solid fuels	Biomass and peat	Liquid fuels
50–100	300	250	300
100–300	200	200	150
> 300	150	150	100

Table 2.3. *SO_2 emission limit values (mg/Nm^3) for large combustion plants, calculated in 6 % O_2. (IED 2010)*

Total rated thermal input (MW)	Coal and lignite and other solid fuels	Biomass	Peat	Liquid fuels
50–100	400	200	300	350
100–300	200	200	300 / 250[a]	250
> 300	150	150	150 / 200[a]	100

[a] in case of fluidized bed combustion

The emission limits presented on the tables 2.2 and 2.3 are considerably tighter than the previous ones. For example, for biomass NO_x limits are tightened approximately 25 % and SO_2 limits are tightened 25 % for large (over 300 MWth) plants. Since emission limits are fuel-specific, for multi-fuel plants, emission limit values are defined as a weighted average of fuels' emission limits according to a used fuel share. (IED 2010) In order to follow the compliance of emission limits, combustion plants (over 100 MW) are obligated to continuous monitoring of NO_x and SO_2 emissions. The monthly average must not exceed emission limits set and daily average need to be below 110 % of those limits. (IED 2010)

3 DATA-DRIVEN MODELING OF BOILER EMISSIONS

Data-driven boiler emissions modeling allows us to gain information about process behaviour and enables the implementation of decision-support systems for both, boiler engineering and operating. Model driven decision support systems employs models for process monitoring and optimization. (MacGregor et al. 2012) In the data-driven modeling the connections between process variables are found from the large amount of process data by utilizing computational intelligence, statistical methods and machine learning algorithms (Solomatine et al. 2008)

Data-driven methods have studied to perform well in CFB emission modeling and it has presented to have certain advantages, especially in real-time monitoring and decision-support applications, mainly because of the online modeling ability. Liukkonen and Y. Hiltunen (2016) highlight that data-driven models can adapt to the prevailing process conditions and they also learn from process history data. In addition to that data-driven models may even be able to predict future states of the process. In the development of data-driven models a detailed and deep understanding of modeled process is not that crucial, than in the programmed computing approach (Krzywanski et al. 2014).

On the other hand some challenges may appear when applying data-driven methods to emission modeling. The key issue is that the performance of model is completely dependent on the quality of the used process data, since model can explain only the information that is presented in the data (Koikkalainen 1994). Therefore data needs to be gathered from sufficient number of separate process states. It is also important to notice that some information, for example fuel properties, may not be available in process measurements. In commercial applications the main challenge is often a long term maintainability of the model, since process conditions typically changes during the plant operation. Because the data that is used for model development may not describe the process behaviour anymore, model need to be trained during the operation. Therefore it is important to have

a balance between model accuracy and a long term maintainability. (Korpela et al. 2017; Martinsuo et al. 2018)

Several approaches have been applied for modeling the emissions of both, biomass and coal combustion, in CFB boiler. These approaches have both limitations and advantages, and the key issue is to find the one that is appropriate, in terms of the structure and complexity, for the application in question. Statistical methods, such as principal component analysis (PCA) and partial least squares (PLS), may be applied to data before developing an emission model, in order to limit the number of needed process variables (MacGregor et al. 2012). Liukkonen, T. Hiltunen et al. (2010) have utilized computational intelligence, more precisely genetic algorithms, to find out delays in the process dynamics. For the CFB emission modeling, both unsupervised and supervised machine learning algorithms, have been presented in the literature. Liukkonen, T. Hiltunen et al. (2011) and M et al. (2011) has used unsupervised algorithms, self-organizing map (SOM) and k-means, to cluster separate processes states in CFB combustion and illustrate NO_x emissions in those states. M et al. (2011) has applied SOM to cluster separate process states form process data to develop NO_x emission monitoring application. Self-organizing maps are popular unsupervised neural networks. Those are known for their ability to visualize multi-dimensional data in a two-dimensional map. Therefore it is often presented to be applied to the decision-supporting systems.

Whereas M et al. has studied a lot of unsupervised learning for CFB emission modeling, Krzywanski et al. has investigated supervised learning to for it and applied multi layer perceptron (MLP) neural network for predicting both NO_x (2014) and SO_2 (2017) emissions in CFB combustion. MLP is a popular statistic method to solve non-linear regression problems. In addition to MLP, there are other network structures applied to the emission regression problem in the literature. One of those is a recurrent neural network (RNN), which is a dynamic network structure. Because of the internal state memory, RNN can process sequences of inputs, unlike MLP (Haykin 2008). This is a great advantage when modeling a highly dynamic process, such as emission formation in CFB combustion. Dynamic models seems to be more accurate than static in the prediction of NO_x emission from the CFB boiler. However, the structure of the recurrent neural network is somewhat complex and therefore, training of RNN requires more computational time and memory than training MLP (Haykin 2008).

3.1 Multilayer Perceptron

Neural networks are systems inspired by the human brain functions. Their potential for solving complex problems has been recognized already back in the mid-1980s when the enormous numbers of algorithms were patented in multiple fields. Back then, the neural networks did not breakthrough in the commercial applications of the process and energy industry. (Koikkalainen 1994) However, at the end of this decade, neural networks have started to gain significant popularity among various fields after a long silent period in the research again. Alongside the traditional neural networks, new, more complex, deep learning algorithms have been developed. This, together with the rapid development of data processing and increased computational power, is considered to be the main enablers for the new rise of neurocomputing (Wu et al. 2018).

A neural network is a generic term that encompasses a wide class of network structures and learning algorithms. Therefore they can be implemented into varying types of problems from industrial processes modeling to image recognition and natural language processing. (Wu et al. 2018). In emission CFB emission modeling applications, neural networks are often applied to solve either regression problem or clustering problem. The emissions prediction problem is a supervised regression problem, which is typically solved by using MLP neural network, which is trained with the backpropagation algorithm. This the most classic combination of neural network and training algorithm. Even though sometimes MLP neural networks may be considered as a complex system, "black box", but they solve a regression problem just by fitting a group non-linear statistical models between input and output parameters (Hastie et al. 2008). MLP trained with backpropagation has been studied to produce an accurate emission prediction. However, the accuracy of the prediction depends strongly on the data quality and chose tuning parameters, such as the number of neurons in the hidden layer. (Krzywanski et al. 2014; M et al. 2011)

3.1.1 Structure

Human brains inspire neural networks and therefore, the base element of the neural network is an information-processing unit called a neuron. Figure 3.1 visualizes a model of a single neuron. Neuron encompasses three elements: A weighted set of synapses, in which each input element x_j of the neuron is multiplied with the specific weight w_{ij}. A sum function for forming a linear combination of v_i of the weighted inputs together with a bias term b_i. An activation function $\sigma()$ for mapping the linear combination v_i to the limited output range. (Haykin 2008)

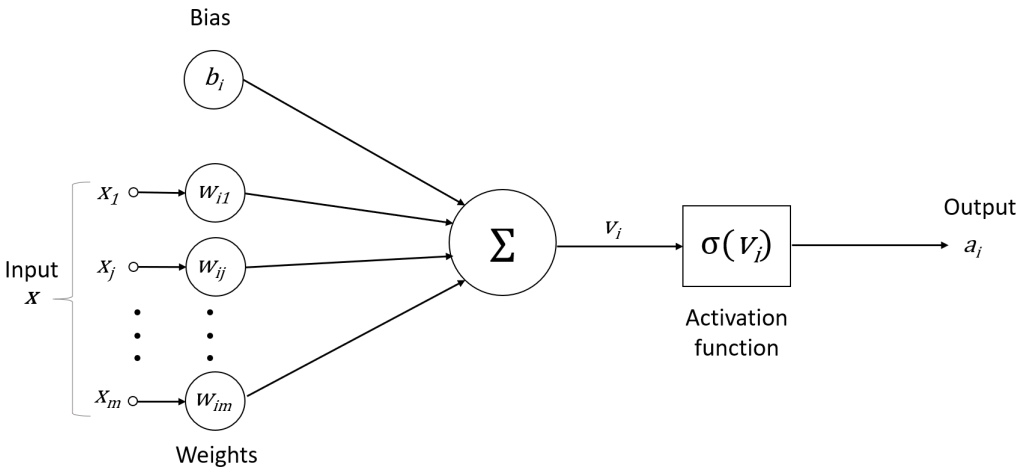


Figure 3.1. *A model of a neuron. Adapted from (Haykin 2008)*

The working principle of a neuron can be expressed in the mathematical form with following two equations:

$$v_i = \sum_{j=1}^{m} w_{ij} x_j + b_i \tag{3.1}$$

where x_1,...,x_m and w_{i1},...,w_{im} and b_i are inputs, weights and a bias of a neuron i, respectively, and v_i is a linear combination of those and

$$a_i = \sigma(v_i) \tag{3.2}$$

where an activation function $\sigma(v_i)$ maps the linear combination to the output a_i. (Haykin 2008)

There are various activation functions studied in the literature and figure 3.2 presents the features for four of those. The first one is an identity function, which assigns the output to be the linear combination itself. It simplifies a neural network to be a linear regression model and therefore, it cannot perform complex problems. The next two ones are the most traditionally used activation functions: sigmoid and hyperbolic tangent. Those are both naturally non-linear functions. The main difference in them is that sigmoid maps the output between zero and one, whereas hyperbolic tangent allows the output also to get negative values. (Hastie et al. 2008) These two to are also often applied to the CFB emission models (Krzywanski et al. 2014; M et al. 2011).

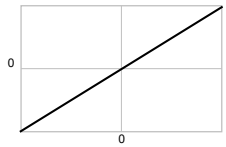
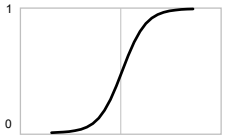
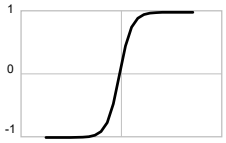
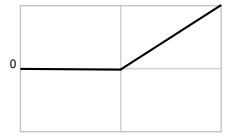
Activation function	Function	Range	Linearity
Identity function	$\sigma(v) = v$ 	[-∞, ∞]	Linear
Sigmoid function	$\sigma(v) = \frac{1}{1+e^{-v}}$ 	[0,1]	Non-linear
Hyperbolic tangent function	$\sigma(v) = \tanh(v)$ 	[-1,1]	Non-linear
Rectified linear unit function	$\sigma(v) = \max(0,v)$ 	[0, ∞]	Non-linear

Figure 3.2. *Activation functions. Data gathered from (Glorot et al. 2011; Hastie et al. 2008)*

The last activation function presented in the figure is a rectified linear unit, which is a relatively new activation function compared to the previous ones since it Glorot et al. have presented it in 2011 for the first time. Ever since, it has been a prevalent activation function, especially in complex networks, since it allows must faster learning in a multilayer network than the traditionally used activation functions sigmoid and hyperbolic tangent. (Glorot et al. 2011; LeCun et al. 2015)

The structure of the neural network consists of half or fully connected layers of neurons. Multilayer perceptron has created a base for the development of the neural networks and it still a commonly used neural network structure, also for emission modeling, event tough Haykin has presented it already back in the 1980s. MLP is a neural network that includes at least one hidden layer of neurons between the input and output layers. Figure 3.3 shows a structure of a fully connected MLP network with four layers: the input layer, the first hidden layer, the second hidden layer and the output layer. As seen from figure 3.3, a multilayer perceptron is a feedforward network, which means that the input information is fed only forward from the input layer layer-by-layer to the output layer. Therefore feedforward networks are static. (Haykin 2008)

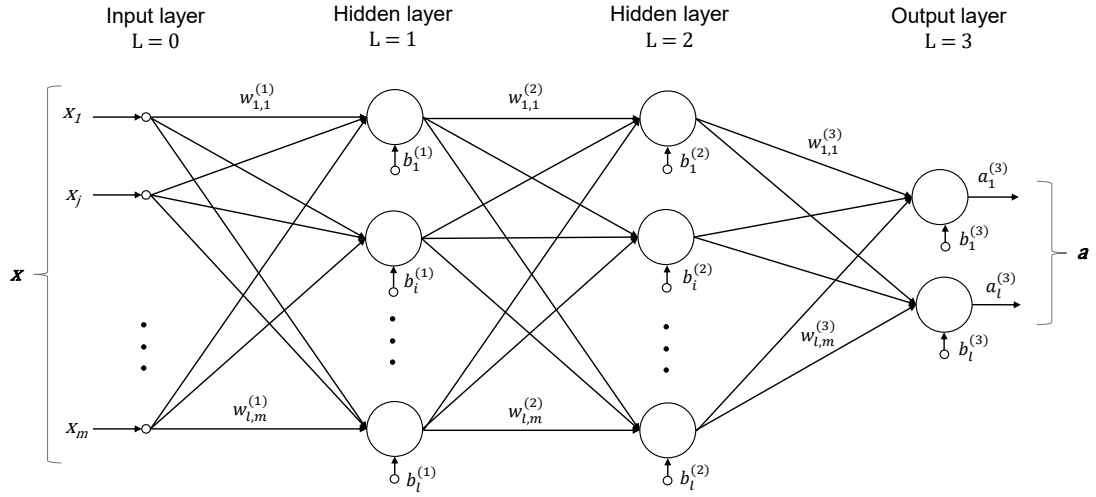


Figure 3.3. *An architectural graph of a MLP with two hidden layers. Adapted from (Haykin 2008).*

For a neuron in the MLP network the equations 3.1 and 3.2 of a single neuron can be now expressed as follows

$$v_i^L = \sum_{j=1}^{m} w_{ij}^L a_j^{L-1} + b_i^L \tag{3.3}$$

$$a_i^L = \sigma^L(v_i^L) \tag{3.4}$$

where i denotes a single neuron in the layer L and inputs for neuron i are the outputs a_j of the last layer $L-1$. For the M layered MLP neural network, the system equations are given in a vector from by

$$\boldsymbol{a}^0 = \boldsymbol{x} \tag{3.5}$$

$$\boldsymbol{a}^L = \boldsymbol{\sigma}^L(W^L \boldsymbol{a}^{L-1} + \boldsymbol{b}^L) \tag{3.6}$$

where $L = 0, 1, ..., M-1$ denotes the layer, $\boldsymbol{x}$ denotes the input vector of the network, $\boldsymbol{a}^L$ denotes the outputs, W^k denotes all the weights, $\boldsymbol{b}^L$ denotes the biases and $\boldsymbol{\sigma}^L$ denotes the activation function of the layer L. (Hagan et al. 1994) It is important to notice that different activation function typically is applied to the hidden layers than to the output layer. The number of hidden layer neurons typically depends on the complexity of the problem to which MLP is applied. For example, Krzywanski et al. (2014) has studied that four-layer [15-9-12-1] MLP structure is accurate when predicting NO_x emissions of coal-fired CFB boiler. Whereas Korpela et al. (2017) has studied that MLP with one hidden layer with two or four neurons performs better than MLP with a hidden layer with six or seven neurons in the prediction of NO_x emissions of the natural gas-fired boiler.

3.1.2 Training

MLP neural networks utilize supervised learning. In supervised learning, the neural network learns to map the associations between an input and an output based on the example input-output pairs, which are referred as training data. The training data consists of the input and output sets $(\boldsymbol{x}_k, \boldsymbol{y}_k)$, in which $\boldsymbol{x}_k$ is the input and $\boldsymbol{y}_k$ is the desired output and

$k = 1, 2, ..., n$ denotes the number of the observation in the data set. The whole training data set of size n is denoted $X = \{(\boldsymbol{x}_1, \boldsymbol{y}_1), ..., (\boldsymbol{x}_n, \boldsymbol{y}_n)\}$. (Hastie et al. 2008) When considering the case of emission prediction, the input $\boldsymbol{x}_k$ could consist of a group of process measurements, with the timestamp k, and output could be the observed emission value at the corresponding timestamp k, for example.

Training of a neural network is an iterative process, in which the goal is to fit the neural network to model the training data, so that the average error between desired output $\boldsymbol{y}_k$ and predicted output $\boldsymbol{a}_k^M$ of the network minimized over the whole training set. This is done by adjusting the weights and biases of the network. Therefore training a neural network can be considered as a complicated optimization problem, where the goal is to minimize the cost function. One of the most typical cost functions is the mean square error. The average of mean squared error over the whole training data set is

$$C = \frac{1}{2n} \sum_{k=1}^{n} (\boldsymbol{a}_k^M - \boldsymbol{y}_k)^2 \tag{3.7}$$

where $\boldsymbol{a}_k^M$ is the output which neural networks predict from the input $\boldsymbol{x}_k$ and $\boldsymbol{y}$ is the desired value for the corresponding input. (Haykin 2008)

Even though training a neural network is a global optimization problem, local optimization methods, such as the Levenberg-Marquardt algorithm (Hansen et al.), conjugate-gradient backpropagation (Moller) and stochastic gradient descent backpropagation (Moller), are often applied for the training. When using the local optimization algorithms it is important to take account that those are fundamentally limited to local solutions, even though that can be avoided by using randomly initialized weights in repeated training, at least with small networks. The performance of these algorithms is also dependent on the problem type. For training, a small network Levenberg-Marquardt algorithm is very efficient, whereas, for the training of a large network, it requires an unreasonable amount of memory. The stochastic gradient descent has an opposite trade of. It's not very fast, but on the other hand, it doesn't require as much computational capacity as Levenberg-Marquardt. Therefore it is more often applied for training large neural networks. (Ilonen et al. 2003)

Stochastic gradient descent is an approximation of the steepest gradient descent. The key idea in the gradient descent based algorithms is to make step changes iteratively to the weights proportional to the negative gradient of the cost function at the current weights until the local minimum is reached. Whereas the gradient descent algorithm calculates

the gradient for the average cost function of the whole training data set (equation 3.7), the stochastic gradient descent approximates the gradient of the cost function only from a randomly selected single training sample or randomly selected mini-batch of the training data set. The cost function is then approximated by

$$\hat{C} = \frac{1}{2}(\boldsymbol{a}_k^M - \boldsymbol{y}_k)^2 \tag{3.8}$$

where the sum is replaced by a squared error of one input-output pair, the stochastic gradient descent algorithm updates the weights as follows.

$$w_{ij} := w_{ij} - \eta \frac{\partial \hat{C}}{\partial w_{ij}} \tag{3.9}$$

where w_{ij} is the j^{th} weight of the i^{th} neuron and η is the learning rate, which defines the step size. (Friedman 2002; Haykin 2008) Because of adding the randomness into the optimization procedure, stochastic gradient descent has studied to enable faster iterations and better accuracy compared to the gradient descent. However, stochastic gradient descent has some trade-off as well. For example, in the large neural networks, the convergence rate might be relatively low. (Friedman 2002; Ilonen et al. 2003)

In the training of MLP networks, typically, backpropagation algorithms are used to compute the gradient of the cost function. Backpropagation has two phases: the forward phase and the backward phase. The forward phase propagates the input $\boldsymbol{x}_k$ forward in the network and calculates the output $\boldsymbol{a}^L$ for the each layer by using the equations 3.5 and 3.6. The backward phase propagates backward layer-by-layer computing the sensitivities for each layer which are defined as follows

$$\delta_i^L = \frac{\partial \hat{C}}{\partial v_i^L} \tag{3.10}$$

where δ_i^L is a sensitivity of the cost function to changes in the linear combination v_i^L in neuron i in layer L.

By using the decencies in the equations 3.3 and 3.4 together with the equation 3.9 it is possible the express the gradient of the cost function with the chain rule

$$\frac{\partial \hat{C}}{\partial w_{ij}^L} = \frac{\partial \hat{C}}{\partial v_i^L} \frac{\partial v_i^L}{\partial w_{ij}^L} \tag{3.11}$$

Where is presented that since the cost function $\hat{C}$ depends on the linear combination v_i^L, which further depends on the weight w_{ij}^L, the cost function $\hat{C}$ depends on the w_{ij}^L as well. Further it can be shown, by applying 3.10 to 3.12 and taking a derivative of 3.3 in respect to w_{ij}^L, that

$$\frac{\partial \hat{C}}{\partial w_{ij}^L} = \delta_i^L a_j^{L-1} \tag{3.12}$$

where the gradient of the cost function for weight w_{ij}^L in the neuron i in the layer L is dependent on the sensitivity δ_i^L of the same neuron and the output of the previous layer. It can be shown that sensitivities have a recurrence relation in the following layers as follows

$$\boldsymbol{\delta}^L = \dot{F}^L(\boldsymbol{v}^L) W^{L+1^T} \boldsymbol{\delta}^{L+1} \tag{3.13}$$

where $\dot{F^L}$ is a diagonal matrix of derivatives of the layer's activation function (of each unit), W^{L+1} is the matrix of the weights of the following layer. Since the backpropagation algorithms begin the propagation from the final layer, the recurrence relation is initialized in it with the equation

$$\boldsymbol{\delta}^M = -\dot{F}^M(\boldsymbol{v}^M)((\boldsymbol{a}_k^M - \boldsymbol{y}_k)) \tag{3.14}$$

which fixes the sensitivity to the error between observed output and predicted output. (Hagan et al. 1994)

When training an MLP neural network, it should take account that its performance is highly sensitive for feature scaling. Differences in the scale across the inputs, may increase the mathematical complexity of training and make it unstable since the scale of the inputs determines an effective scaling of the weights in the first network layer. (Hastie et al. 2008) To avoid this, the input data is typically standardized.

The standardized value for a sample x is calculated as

$$x_{standardized} = \frac{x - \mu}{\sigma} \tag{3.15}$$

where μ is the mean of the training samples and σ is the standard deviation of the training samples. As a result of the standardization each input have zero mean and unit variance. (Pedregosa et al. 2011)

3.1.3 Evaluation Metrics

Evaluation metrics evaluate model performance. In the regression problem, this means the evaluation of the model's ability to match the observed value to the corresponding predicted one. Typical evaluation metrics are R^2 (also referred to as the coefficient of determination and MAE (mean absolute error).

R^2 evaluates the goodness of the model fit to a particular data set and how well the model will predict the unseen data samples. It is dimensionless statistics, which indicates the proportion of the variance of the output variable that input variables explain. One definition of it is

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{3.16}$$

where y_i is the observed value, $\hat{y}_i$ is the corresponding predicted value and $\bar{y}$ is the mean of the observed data. (Kvalseth 1985) Values of R^2 typically ranges from zero to one, however, it is also possible for it to get negative values, if the model performance is really poor. The highest score for R^2 is one, which implies that model is perfect where, and therefore a high R^2 scores are preferable. (Alexander et al. 2015)

Mean absolute error is a simple statistic method for the regression evaluation and they are defined as follows

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|. \tag{3.17}$$

Opposite to R^2 for MAE error small values are preferable. (Kvalseth 1985)

4 MATERIALS AND METHODS

In this study, data-driven models of CFB boiler NO_x, SO_2 and CO emission were implemented by using MLP neural networks. All these models were trained individually and besides those, an MLP model for operational costs of emission control was developed. These models were implemented so that they could be exploited further in the proposed cloud-based air emission advisor application. The key idea in the air emission advisor application concept is to support the operator's decision making by suggesting changes in certain process variables. These advice help operators to achieve daily emission targets in varying process conditions with minimized costs. The proposed air emission advisor application concept bases on the NO_x emission advisor application developed in Valmet.

Figure 4.1 presents generalized schematics of the proposed emission advisor concept. The figure shows that models are developed in the offline environment, whereas the application would run online and deploy the trained models for the multivariate optimization of process conditions. As emphasized in the figure with the dashed line, this work focuses on the offline development of models, which encompasses phases, such as data gathering and preprocessing, model inputs and outputs selecting together with MLP training and validation. These phases are described more detailed further in this chapter. Also, the variables, called operational variables that would be optimized in the advisor application are selected in this work, since it is essential to evaluate the performance of MLP models not only with the prediction accuracy but also with predicted correlations between emissions and operational parameters.

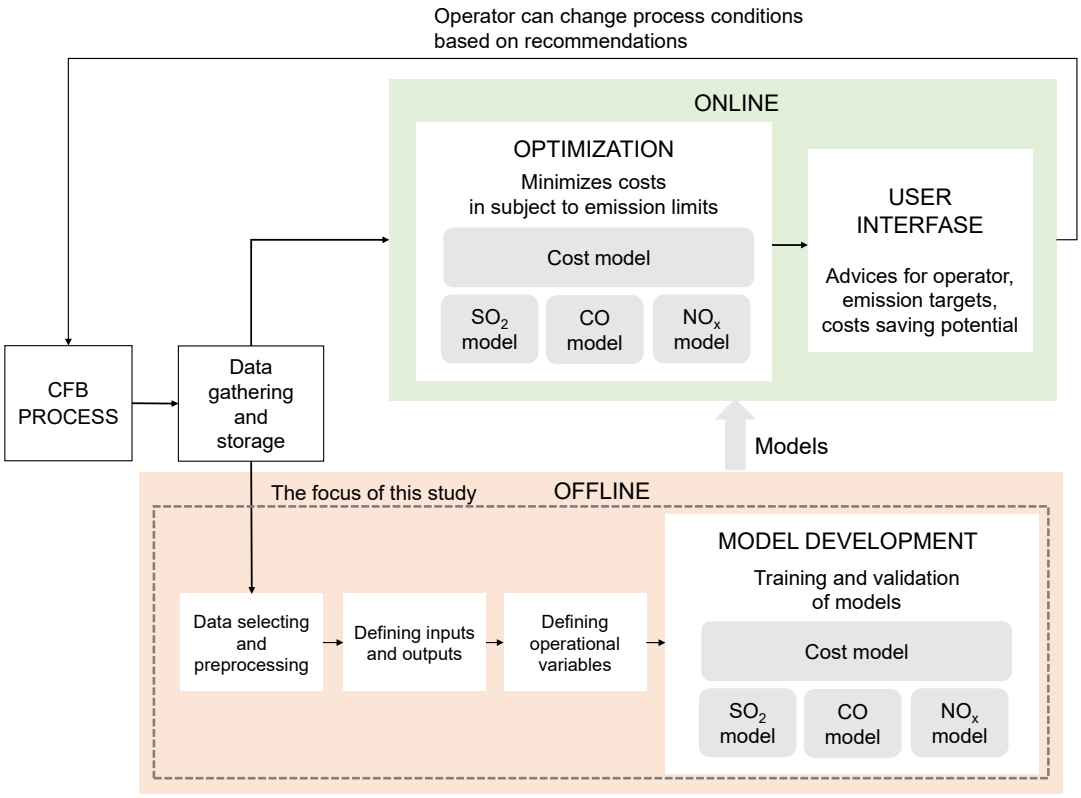


Figure 4.1. *Generalised schematics of the proposed emission advisor concept*

As the figure 4.1 presents both, data-driven modeling in the offline environment and the online application rely on the CFB process data. The history data of the process is stored to some data storage, to the cloud in this case, and then utilized in the models' development phase. For the online application, the real-time data stream of the same process parameters is used. The case in this study relies on the process data gathered from a reference plant with the CFB boiler. The reference plant is a CHP plant, which produces district heat for surrounding regions and process steam for local industry combined with electricity. The boiler in the plant is a circulating fluidized bed boiler, equipped with two cyclones, which Valmet has supplied. It is a multi-fuel boiler, which uses the biomass-coal mix as it's primary fuel.

The emission control in the plant is based on both combustion conditions optimization and additives feeding. The formation of nitrogen oxides in the furnace is limited to the air staging. In the reference boiler, combustion air is divided into the primary and secondary airs and the secondary air consists of the upper and lower secondary airs, which are divided even further to the sections called must and rest airs. Names must and rest imply their necessity for the complete combustion. The flue gas NO_x emissions are reduced with the SNCR method, in which ammonia is injected in the flue gas either from the upper

part of the furnace or from the cyclones inlets, depending on the process state. Sulfur emissions are reduced with limestone feed, which is typical for CFB boilers. Limestone is fed to the lower part of the furnace from which it mixes into the bed material.

4.1 Data Selecting and Reprocessing

Process history data from the Valmet cloud database was used for models training, validation and testing. Data has been gathered during the real plant operation. Initially, it has been generated and stored in the local automation system of the reference plant, from which it has been transferred and stored to the cloud for further use. Only the process data available online was chosen to be used in modeling since models should be capable of real-time modeling in the emission advisor application. History data used for the models' development consisted of 148 days of continuous, one minute averaged process data from the reference plant normal operation. The data was chosen to be rather long to maximize the diversity of operating points, for example, varying boiler loads and fuel mixes, in data. The set contained data of over 50 process variables from which inputs and outputs for each model are defined in the following subsection.

The history data file was downloaded from the cloud database and the model development was done in the offline environment by using the Python programming language. Data was read and stored in the data frames by using the functions from the open-source library called Pandas. The data was preprocessed before used for models training. The first step was to make sure that all data was consistent. For example, some of the data appeared with varying decimal separator and time format. After that, the missing values were removed and data was filtered from the boiler shut down periods. In addition to these, the noisy emission measurements, especially SO_2, was tested to filter with two different methods of rolling mean and wavelet transform. The wavelet filtering was done by using an open-source library pywt by Lee et al. (2019). However, it was found out to be rather tricky to define which information is just noise in data an which is caused due to some process change. Therefore filtering was decided to drop out from the data preprocessing phase.

During the data preprocessing, it was also taken into account that emission formation and reduction are highly dynamic processes, whereas the MLP is a static model. This may lead to the situation in which correlations between the process variables and emissions

do not describe the real process behavior since the change in the process shows with a delay in the flue gas measurements. For example, if ammonia is injected into the furnace due to the high NO_x levels at the time t and at the same time t high NO_x emissions are measured from the stack. Then MLP learns that ammonia injection implies high NO_x emissions, even though the correlation should be opposite. To avoid that, all emission measurements were delayed with two minutes. The most suitable delay was iteratively selected with simple model training and validation. All emission measurements were sifted with the same delay and the delay was assumed to be a constant for the entire data set. Finally, the whole data set was standardized as described in the equation 3.15. The standardization was implemented with the preprocessing module available in the scikit-learn library.

The preprocessed data set was then divided into subsets for training and validation of the models, as visualized in the upper part of the figure 4.2. The training and validation data sets cover 148 days of one minute averaged process data.

Figure 4.2. *Data was split into the training and validation data sets.*

The training and validation data were divided further into four folds for cross-validation as visualized in the lower part of the figure 4.2. Those folds were selected with a time-based selection so that in each fold training period is followed by a five days validation period. Successive training sets are subsets of those subsets of which come before them. Also, each validation period has a larger timestamp than the one in the previous fold. These four folds were used for model training and validation when defining the suitable model

parameters. Since folds were chosen with a time-based selection, it is very likely that process conditions changes in between the data sets.

A closer look at the training and validation data was taken to get a rough overview of prevailing process conditions ion each fold since those might influence models performance. Table 4.1 presents just very generalised classification of prevailing conditions for coal share, boiler load, NO_x, SO_2 and CO conditions in each fold.

Table 4.1. *Training and validation data sets*

Fold	Coal share	Boiler load	SO_2	NO_x	CO
1	Varying	High	High	Varying	Low
2	High	High	Varying	Low	Low
3	Varying	Varying	Varying	Varying	Low
4	Low	Low	Low	High	Low

The table reveals that process conditions seem to be varying a lot in between the folds. For example, the boiler load is high in the first two folds, then changing in the next fold from high to low and remains low in the fourth fold. There are also differences in NO_x and SO_2 emission levels between the folds, whereas carbon monoxide emissions levels remain low in each fold.

4.2 Model Input and Output Variables

Next suitable input and output variables for each model were chosen. First, potential inputs were selected based on the knowledge gained from the literature review and Valmet process specialist. It was concluded that at least variables from following, fuel, air feed and additives feed, bed and chamber conditions and flue gas should be used. Air, feed and additives feed, bed and chamber conditions and flue gas. Further, measurements of the main steam flow and shoot blowing flow were needed have. In addition to the variables describing process state and emission conditions, ones were needed to define to model operational costs. Input variables for the cost model were selected from the group of emission models inputs. The final list of models variables was selected with a simple regression model by using leave one out method. The list of the input and output variables for each model is presented in the table 4.2.

Table 4.2. *Variables used in the models.*

Variable	NO_x	SO_2	CO	COST
Fuel feed				
Coal share	X	X	X	
Air feed				
Primary air ratio	X	X	X	X
Additives feed				
Ammonia flow	X			X
Limestone flow	X	X		X
Ammonia feeding mode	X			
Bed and chamber conditions				
Bed temperature	X	X	X	
Dipleg temperature	X	X	X	
Recirculation gas flow	X	X	X	X
Steam flow				
Main steam flow	X		X	X
Shoot blowing steam flow			X	
Flue gas				
Flue gas temperature after cyclone 1	X		X	
Flue gas temperature after cyclone 2	X		X	
Flue gas O_2	X	X	X	X
Flue gas ammonia slip	X	X		
Flue gas NO_x content	X			
Flue gas SO_2 content		X		
Flue gas CO content			X	
Costs				
Emission control cost				X

As the table shows, most of the models inputs are process measurements. In this case, most of them were available from the plant automation system, however, some of the model inputs were calculated based on process measurement or other signals gathered from the process. These input variables are coal share, which describes the share of coal in the fuel mix, primary air ratio, which describes the completeness of air staging and ammonia feeding mode, which describes the location of the ammonia injection. These variables are combined information from several process signals to reduce the number of inputs and, thus, the complexity of the models. The coal share is calculated as described in the equation below,

$$Coal\,share = \frac{Coal\,load}{Total\,fuel\,load}. \tag{4.1}$$

Where the total fuel load is a sum of coal load and biomass load, in this case, onwards, the primary air ratio is defined as follows.

$$PAR = \frac{Primary\,air\,flow}{Total\,combustion\,air\,flow} \tag{4.2}$$

Where the total combustion airflow is a sum of secondary, loop seal and leakage air flows. Ammonia feeding mode describes where the ammonia is injected based on the position of ammonia injection valves. Ammonia feeding mode gets integer values from one to three as follows: ’1’ for ammonia injection to the furnace, ’2’ for ammonia injection to the cyclone and ’3’ for no ammonia injection.

Outputs for the emission models were rather straightforward to define, since measurements of NO_x, SO_2 and CO are each an output of the corresponding emission model. In contrast, one for the cost model required further calculations. The output for the cost model, emission control cost, was calculated as described in the equation 2.4. The table 4.3 presents prices that were substituted together with the process measurements to the equations 2.1, 2.2 and 2.3. When defining the losses in the boiler efficiency, it needed to take account in this case that due to local legislation, the coal is priced differently depending, whether it is used for heat or electricity supply.

Table 4.3. *Constants used in the cost calculation*

Consumable	Price
Additives	
Ammonia price	250 €/t
Limestone price	100 €/t
Power consumption	
Electricity price	40 €/MWh
Boiler efficiency losses	
Coal price, heat production	50 €/MWh
Coal price, electricity production	20 €/MWh
Biomass price	30 €/MWh

It should be noted that the primary purpose of the cost model not provide the absolute cost of emission control but catch the change that changes in the operational variables cause to the costs. Therefore certain factors, like emission control is not the only cause for fans' power consumption, were not taken account in cost calculations.

4.3 Operational Variables

The main purpose of the emission models is not to be a soft sensor predicting emissions since there are stack measurements of the emissions available. Instead of that, the key idea of emission modeling is to model correlations between certain operational variables and emissions. These operational variables are selected from the process variables with two criteria: They are inputs for at least one emission model and the operator can change them during a boiler operation at least within certain limitations. Based on these criteria, operational variables were chosen individually for each emission model. Those are presented in the table 4.4.

Table 4.4. *Operational variables*

Variable	NO_x	SO_2	CO
Flue gas O_2	X	X	X
Primary air ratio	X	X	
Bed temperature	X	X	X
Recirculation gas flow	X	X	
Ammonia flow	X		
Limestone flow	X	X	

The first two operational variables indicate the air excess in combustion. Those are all typically controlled with feedback control, and operators can often change the control strategy with manually set factors. High oxygen content in flue gas indicates high air excess during the combustion. In the reference boiler flue gas, O_2 content varies typically from 2 % to 6 % and it is controlled with feedback control, which aims to follow a boiler load-dependent set point. Primary air ratio asses the completeness of air staging. In the reference case, it gets values between 0.4 and 0.7, where high value assigns a low air staging rate and wise versa. Air staging, and thereby the primary air ratio, is also controlled with feedback control. The combustion air share to the primary and secondary airs is dependent on the boiler load and the fuel share. Further, at the low boiler loads, primary air ratio is limited by the bed height, since it is responsible for the bed fluidization.

It is presented in the literature that bed temperature correlates strongly with every three emissions. It typically varies around 850 ℃, and recirculation gas flow controls it. Also, it seems that flue gas recirculation may reduce both SO_2 and NO_x emissions. In this case, recirculation flow can vary from 0 to 14 Nm^3/s and it does not participate in bed material fluidization.

The primary purpose of the limestone feed is to reduce SO_2 emissions. However, it has an impact on NO_x emission levels as well, whereby the ammonia flow is the only operational variable correlating with one emission, NO_x. The additive feeds are both controlled with the control strategy that combines both the feedforward and feedback control. Naturally, the limestone flow control is based on the SO_2 measurement and ammonia low is then controlled based on measurement.

4.4 Neural Networks Training and Validation

The preprocessed data set was used in the supervised training of MLP neural network models. In total, three emission MLP models and one cost MLP model were trained by using a free, open-source library, scikit-learn, for the Python programming language. Scikit-learn library features various machine learning and data preprocessing and algorithms for data analysis and its widely used for solving simple machine learning tasks. (Pedregosa et al. 2011) This study uses the MLP regressor from the neural network module for the training of the neural network. In addition to these, several functions from the metrics module were used in the models' validation.

Figure 4.3 percents a flow chart of each emission model training and validation process, which reveals that selecting the most suitable model architecture and training parameters was an iterative process. First, initial parameters for model and training were set. Then the MLP model was trained with four-fold cross-validation in which a model for each four-fold, presented in figure 4.2, were trained with corresponding training data set. After that, the trained models were evaluated with the unseen and independent validation data sets. The final evaluation of model performance was then done based on the average of the performance of each four-fold. The cross-validation was used not only to avoid model overfitting but also to examine the robustness of model performance in varying process conditions.

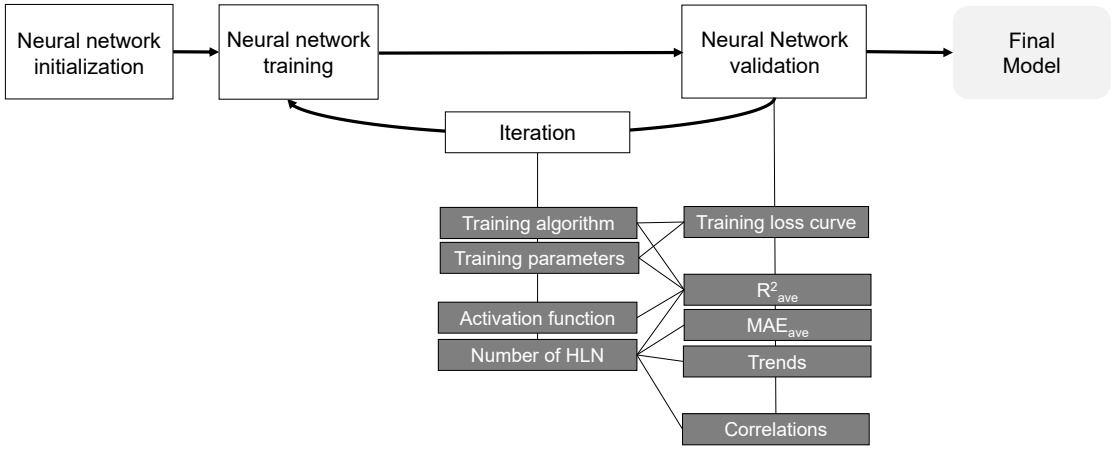


Figure 4.3. *Selecting the final parameters for the emission models was an iterative process.*

As the figure shows 4.3, several parameters were selected during the iterative training and validation process. It also reveals that the parameters were validated with varying validation methods. First, it was examined which the stochastic gradient-based or quasi-Newton based algorithm is more suitable for model training. Then the parameters of the selected training algorithm were adjusted to ensure the training convergence. It was also examined whether the solver can find out a local minimum, providing an acceptable model performance. The goodness of the training algorithms and training parameters were mainly evaluated with the training loss curve. Also, R_2 score was calculated to determine the prediction accuracy reached with different training setups. It's good to emphasize that even though the performance of training algorithms and parameters were tested for different MLP structures, the goal was to find out generally suitable training parameters for the data set rather than finding the best combination of all parameters. A similar principle was followed when choosing an appropriate activation function.

When a suitable training algorithm, it's parameters and an activation function were determined, the best MLP architecture was begun to seek iteratively. First, the number of neurons in the first hidden layer was changed in between one and ten. Then it was examined whether adding a second hidden layer increase the model prediction accuracy. The prediction accuracy of each model architecture was evaluated with evaluation metrics, MAE error and R_2 score, preferring high R_2 and low MAE values. Those evaluation metrics were typically examined as an average of each fold in cross-validation. However, it was taken to account that folds included data from varying process conditions, and therefore, models may perform systematically better for some set than the other. That's why the R_2 scores were also compared between the data folds. Also, if the model performance seems very poor some fold, it is possible to validate the model structure without that one.

The performance model structure providing predictions with the best MAE and R_2 results were examined more closely. These models were compared by plotting trends of those predictions and compared those to the measured emission levels. Also, trends of input parameters were examined to find out whether the behavior of some input parameters may explain the behavior of prediction error. In addition to the prediction accuracy, the correlation between predicted emissions and controllable variables were examined visually. When developing emission models for emission monitoring application, described in figure 4.1, it is more important than models can model the correlations to the controllable variables with an appropriate level, rather than produce accurate predictions in the

absolute level. However, it was challenging to find a statistical method to evaluate this. Therefore models were assessed by plotting correlation between an operational variable and predicted emission and then comparing that to the correlations presented in the literature. Based on all these validation methods, the final emission model structure was selected for each three models.

5 RESULTS AND DISCUSSION

5.1 Emission Models Performance

The performance of emission models were evaluated in four-fold cross validation. The validation scores were compared between model structures. Furthermore, models' ability to correlated with operational variable in similar way than the literature presents were evaluated. Based on the evaluations the most suitable model parameters were selected individually for each emission model. All emission models were decided to train with the same training algorithm and activation function, though.

To find the most suitable training algorithm the stochastic gradient descent (sgd) based training methods were compared to quasi-Newton one. Generally, sgd based methods seemed to converge faster and provided better validation scores than the quasi-Newton one. Therefore, an sgd based solver was selected to be used. Thus, it was noticed that the solver tended to converge occasionally to a local minimum, which led to poor validation scores. To avoid that, a training parameter, called beta, was adjusted individually for each model. In addition to the training algorithm, the most suitable activation function was found out to be used in each emission model. Based on the R^2 scores ReLu was chosen, even though more traditional activation functions, like hyperbolic tangent sigmoid are often suggested for emission modeling in the literature (Krzywański et al. 2017).

5.1.1 SO_2 Model

The most suitable value for training parameter beta and the best model structure was attempted to find out by using four-fold cross-validation. Four validation runs were executed with the data folds presented in figure 4.2. In the cross-validation runs, it was found out that the best validation scores for R^2 were achieved when the training parameter beta got value 0.1. Therefore, that was used in training when examined which model structure provides the best prediction accuracy.

Figure 5.1 presents the effect of the number of hidden layer neurons (HLN) to validation scores for the evaluation metrics R^2 and MAE. When examining the results in figure 5.1, it should be noted that in this case, the validation sets were selected with a time-based selection, not randomly like usually in the cross-validation. Because of that, the validation sets have data from very different process conditions (see figure 4.1), which may decrease the statistical robustness of the cross-validation and cause notable variation of results between the validation runs.

Figure shows that the model with three HLN provides the best validation score (0.64) for R^2. It is encouraging that the R^2 score is just above 0.60. It implies that the model can predict SO2 emissions with an admissible accuracy from unseen data. What is surprising in the figure is that the best validation score for MAE error (13.8 mg/Nm^3) was not provided with the same model. Instead of that, the lowest MAE error was given by the model with seven HLN. It may imply that the model with seven HLN provides predictions with smaller biased error than the one with 3 HLN, whereas it fails to predict changes in emission levels as accurately as the model with 3 HLN.

When comparing the SO_2 model prediction accuracy reached in this study to the one Krzywanski et al. (2014) has presented for SO_2 MLP model there is a major difference in models performance. Krzywanski et al. (2014) presents R^2 scores as high as 0.98 for generalized SO_2 model, which is a significantly higher score compared to the one reached in this study, 0.64. However, differences in model inputs can at least partly explain with the differences in model inputs. In essence, Krzywanski et al. (2014) has used variables like fuel sulfur content, and fuel particle size as an input, whereas the only variable indicating fuel properties was fuel coal ratio. Also, Krzywanski et al. (2014) does not mention whether R^2 scores are achieved with training data or unseen validation data.

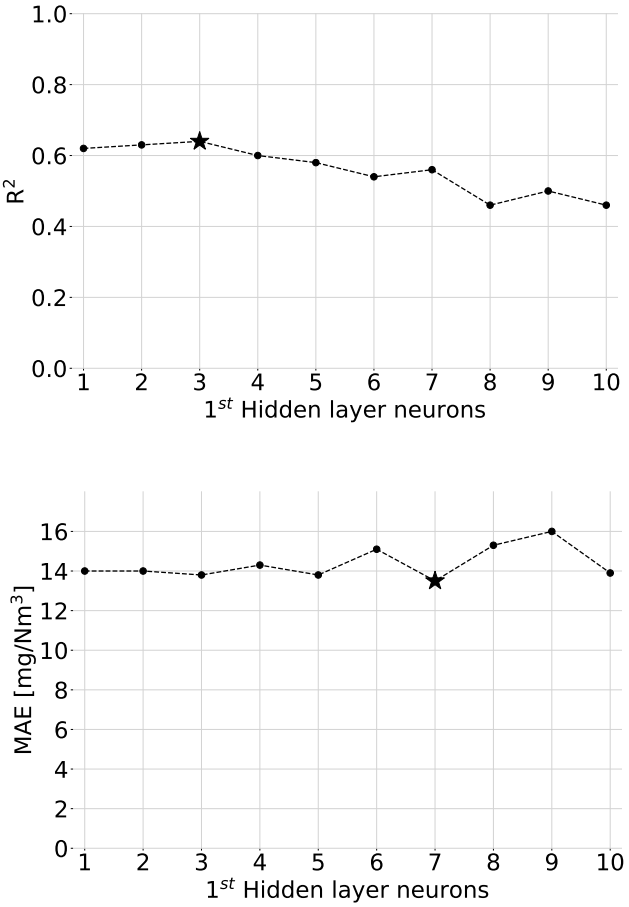


Figure 5.1. *The effect of the number of neurons on validation scores in four-fold cross validation.*

Based on the validation scores, the prediction accuracy of the two best models was examined more closely. These model structures were three HLN (8-3-1) and seven HLN (8-7-1). Figure 5.1.1 presents 12 hours trends of measured and predicted SO2 in each four validation run. The first graph is from validation run one, and the last one is from validation run four. These trends give an overview of how do predicted emissions behave in comparison to each other and the measured ones in each validation run. However, it should be noted that these periods are only a short take from each validation run. Therefore that only supports the decision making, and definitive conclusions can not be made exclusively based on those.

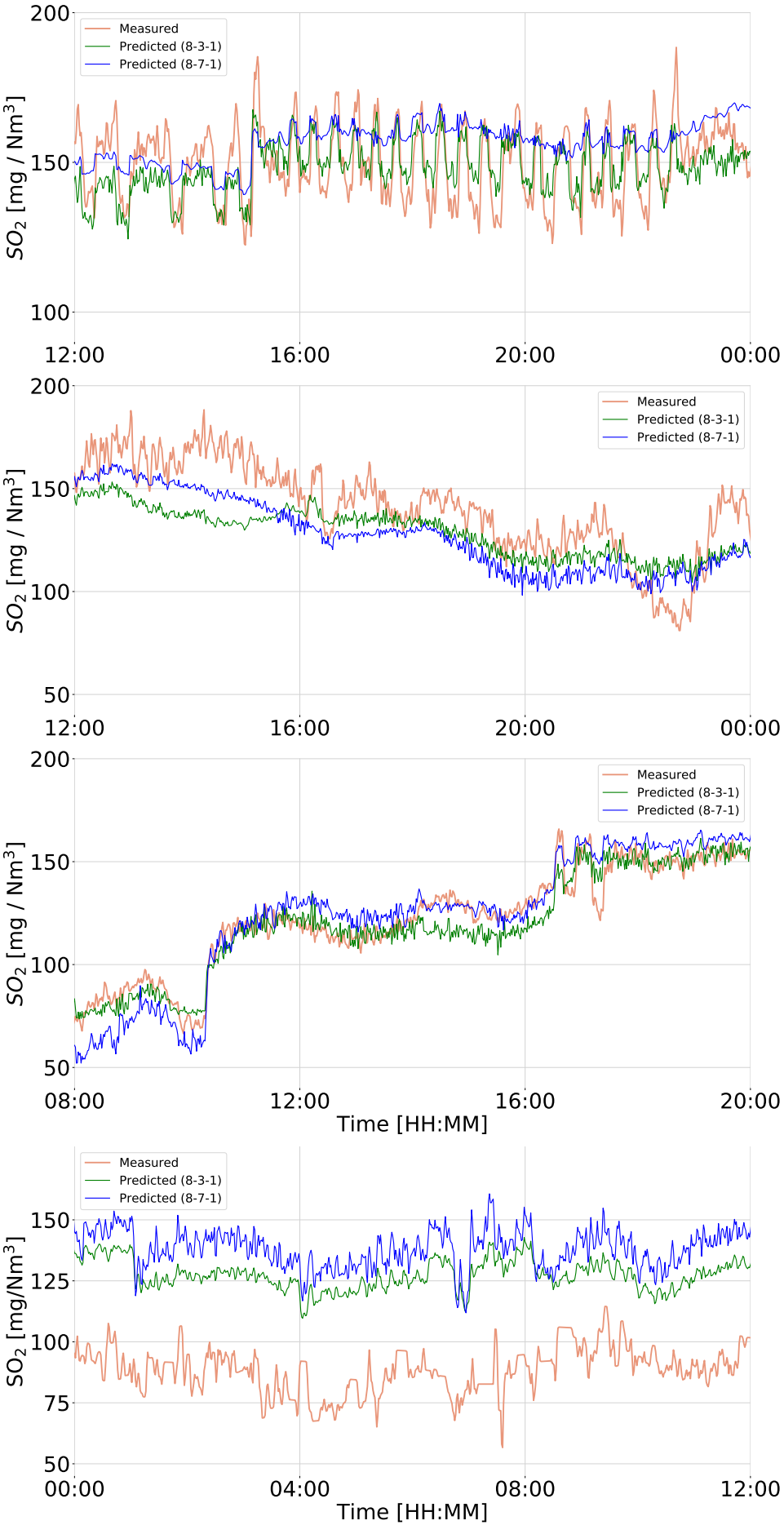


Figure 5.2. *Trends of measured and predicted SO_2 emissions. Predictions are provided with the models 8-3-1 and 8-7-1.*

The first graph presents two predictions and measurements of SO_2 emissions in the validation run 1. It shows that measured values are varying about 30 mg/Nm^3 with quite regular intervals, which are likely caused by the limestone feedforward control. It seems that both models can estimate the general pattern of measured values. Thus, the model 8-3-1 can predict the amplitude of measured emission significantly better than the model 8-7-1 as the validation scores for R_2 implied. When taking a closer look at the input data, it was noticed limestone feed gets on and off with the corresponding interval in the same period.

From the next graph, validation run 2, it is quite challenging to conclude which model performs better. Both predictions seem to follow well slow change in the emission level while they struggle to follow irregular fluctuations in the sulfur emissions. One reason for this kind of performance may be the fact that model inputs do not include information of fuel properties, for example, sulfur content, and therefore models can not predict small changes in SO_2 emissions.

Neither the third graph, validation run 3, reveals significant differences between the models' prediction accuracy since both predictions follow the general pattern of the measurement. However, it seems that model 8-3-1 provides a smaller biased error than the 8-7-1. The substantial step-change in emission measurement may be caused by the sudden growth in fuel coal share. Since the coal share is an input for the model, it is natural that models can predict that change.

In the last graph, validation run 4, both models seem to fail to predict emissions. However, it appears that model 8-7-1 can represent the shape of measured emissions slightly better that model 8-3-1. Both predictions occur with a significant bias error, even over 40 mg/Nm^3. The biased error implies that there are some conditions in the validation data that are not occurring in the training data. For example, fuel sulfur content may have changed. Also, the sulfur level is remarkably lower and in the validation run 4 than in the other ones. Therefore limestone feed is rarely used during validation run 4, which might explain poor prediction accuracy. Since most of the inputs, such as flue gas oxygen content, PAR and bed temperature, correlate to sulfur removal efficiency, not its generation, there are only a few inputs explaining the behavior of emissions when limestone feed is not used.

As a summary, both models succeed to predict emissions with an admissible accuracy in the first three validation run. In the fourth validation, run models seemed to perform

rather poorly. Actually, the behavior of the trends presented in figure indicates that the ability of models to predict changes in the data is might be even more dependent on the process conditions than the model structure. Therefore the major difference in model prediction accuracy between the validation runs may be explained by the fact that the process conditions differ significantly between those as presented in the table. 4.1. In this case, it seems that models provide more accurate predictions in the process conditions where SO_2 emissions are high than in the ones with low emissions. When considering the end use of the emission model, this is an encouraging result since models are most needed in the high emissions.

In this case, model performance should not only be evaluated with the prediction accuracy, but also its ability to model correlations between emission and operational variables. Correlations were studied as follows: First emissions were predicted in an operating point occurring in input data, called base point. Then the value of the operating variable was changed while other inputs were kept constant. That was repeated for randomly selected operation points, four of which are presented in figure 5.3 for models 8-3-1 and 8-7-1. Only process conditions where the limestone feed is on were examined since the operational variables correlate with SO_2 removal efficiency, not with SO_2 generation.

To get an overview of how models can predict correlations between SO_2 emissions and each operational variable, out clouding ammonia feed, in different operating points, the slope of modeled lines were examined more closely. Since there is not process data available describing all predicted states, the modeled slopes were compared to the generalized correlations presented in the literature (see figure 2.9 in chapter 2). The generalised correlations are also represented in figure 5.3.

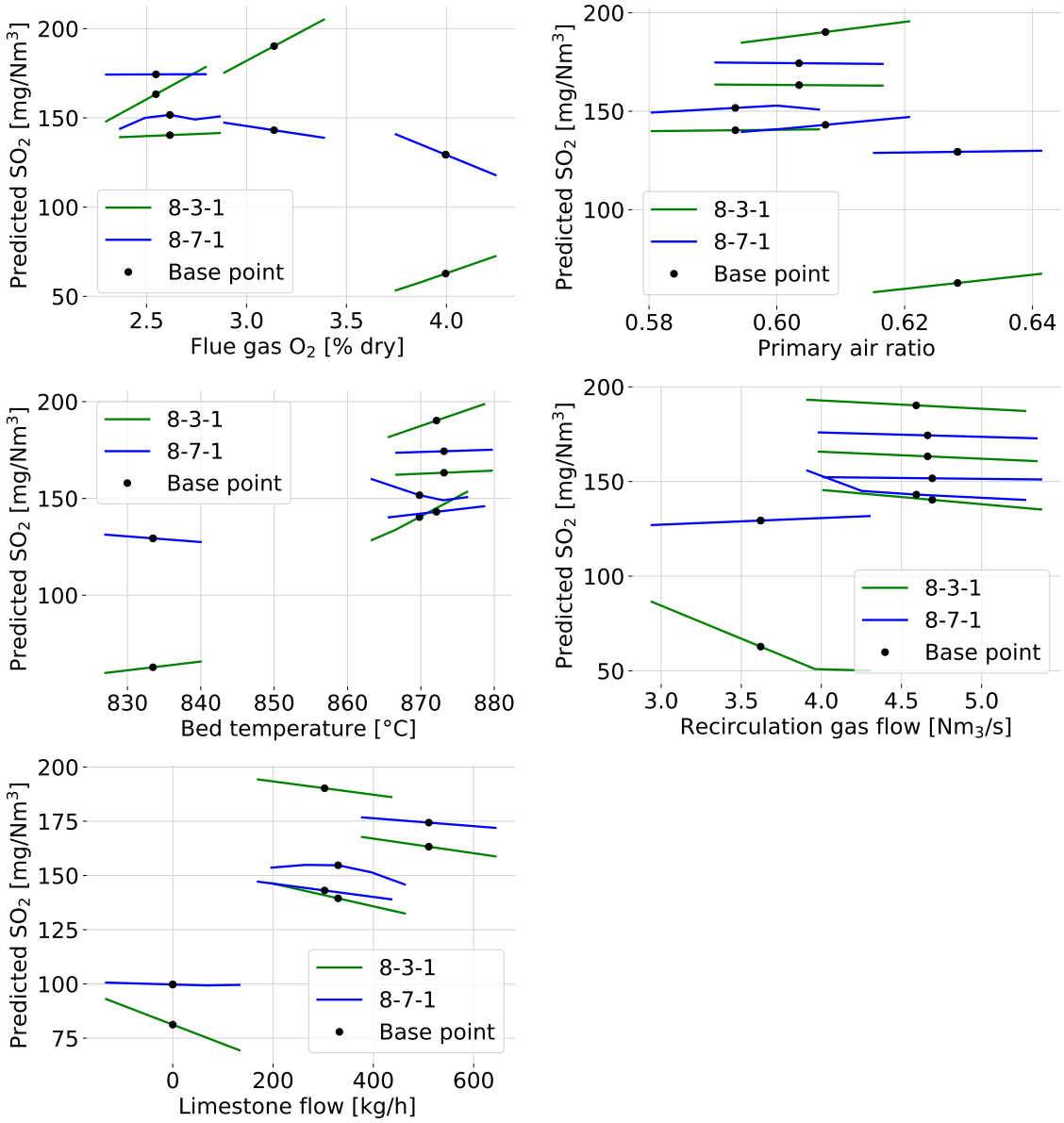


Figure 5.3. *Modeled correlations between SO_2 emission and operational variables. Slope of each line shows modeled correlation around an operation point (base point).*

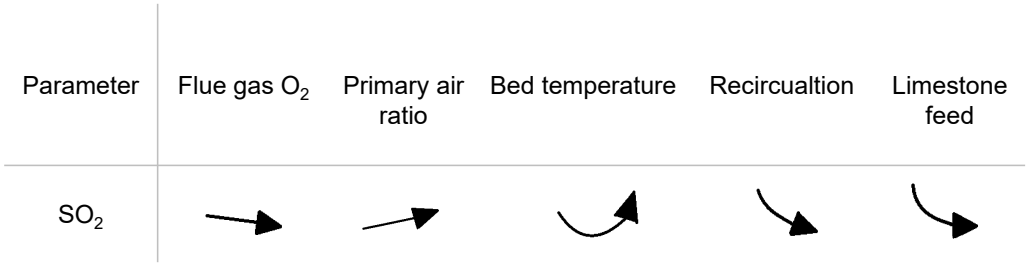


Figure 5.4. *Generalized correlations presented in the literature.*

The first graph shows modeled correlation between SO_2 emissions and flue gas O_2 content. In all four operating points, it seems that the model with 3 HLN predicts an increase in air excess to improve the SO_2 removal efficiency, whereas the model with 7 HLN gives the opposite result. When comparing these observations with the ones presented in the literature, it seems that the model with 7 HLN performs remarkably better. The second graph presents the modeled effect of air staging to which both models provide lines with positive slope: An increase in primary air ratio causes a small rise in emission levels, which correlates to observations presented in the literature.

In the third graph, the predicted correlations between SO_2 and bed temperature are presented. Both models predict that in high temperature levels, around 870 °C, a rise in temperature will increase the emissions. On the other hand, at a lower temperature, about 830 °C, the slope of prediction 8-7-1 is negative, and prediction 8-3-1 is positive. When comparing these results to the ones presented in figure 5.4, it seems that model 8-7-1 follow the parabola shape slightly better, however, it's difficult to draw a definite conclusion of that. When taking a look at the fourth graph, the model dominant slope direction is negative. This indicates that flue gas circulation decreases predicted emission, as presented in the literature, too.

The last graph visualizes the modeled correlation between SO_2 emission and limestone flow. The figure indicates that both models predict limestone addition to reduce emissions with somewhat linear correlation, which was an expected result based on figure 5.3. To conclude, models, seem to provide correlations that are in line with the ones presented in the literature rather well. The only correlation that is not aligned with the research is the flue gas O_2 correlation provided with model 8-3-1.

5.1.2 NO_x Model

Four-fold cross-validation was also used when seeking the most suitable parameters for training and hidden layer number for NO_x model. These validation runs were executed with the data folds presented in figure 4.2. The best validation scores for the R^2 metric resulted when the training parameter beta was 0.9. That value for the beta was used when examining model validation scores in different model structures.

Figure 5.5 shows how does the number of hidden layer neurons (HLN) affect validation scores for the evaluation metrics R^2 and MAE. It reveals that the model with nine HLN achieves the best validation scores for both evaluation metrics. The R^2 score for model 12-8-1 is 0.53, whereas the second highest R^2 score is 0.49, provided by model 12-3-1. The MAE error for model 12-8-1 is 9.1 mg/Nm^3 and 10.6 mg/Nm^3 for model 12-3-1. It is a bit disappointing result that there is not a model that can reach the validation score for R^2 over 0.60.

In this study the model validation scores indicated that adding the second hidden layer to the model architecture does not improve the model prediction accuracy where as Krzywański et al. (2017) present a study where the highest R^2 scores, 0.94, were provided with the (15-9-12-1) NO_x MLP model. Thus, it is difficult to draw an definitive conclusion whether more complex model structure would help to achieve better for the NO_x model in this study. In essence, Krzywański et al. (2017) has used laboratory analysis of fuel properties as an input, whereas the only variable indicating fuel properties was fuel coal ratio. Also, Krzywański et al. (2017) coal combustion in CFB boiler, in which the NO_x formation is more simple than in biomass or co-combustion.

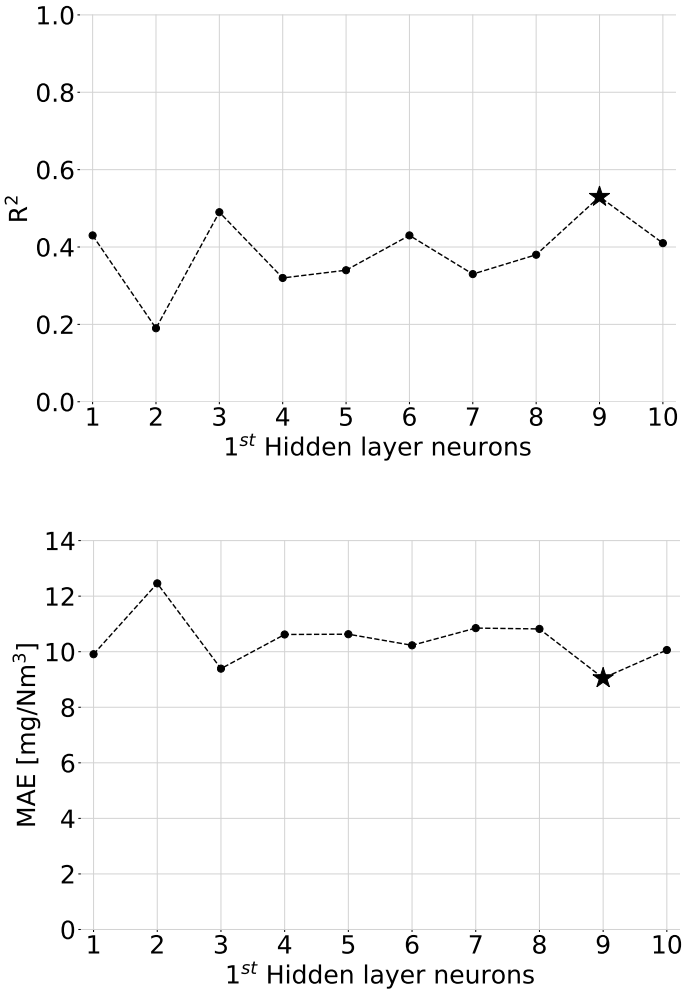


Figure 5.5. *The effect of the number of neurons on validation scores in four-fold cross validation.*

Based on the validation scores, the prediction accuracy of two models with the highest validation score for R^2 was examined more closely. These model structures were eight HLN (12-8-1) and three HLN (12-3-1). Figure 5.6 presents 12 hours trends of measured and predicted NO_x in each four validation run. The first graph is from validation run 1 and the last one is from validation run 4. These trends give an overview of how do predicted emissions behave in comparison to each other and the measured ones in each validation run. However, it should be noted that these periods are only a short take from each validation run and, therefore, that only support the decision making and definitive conclusions can not be made only based on those.

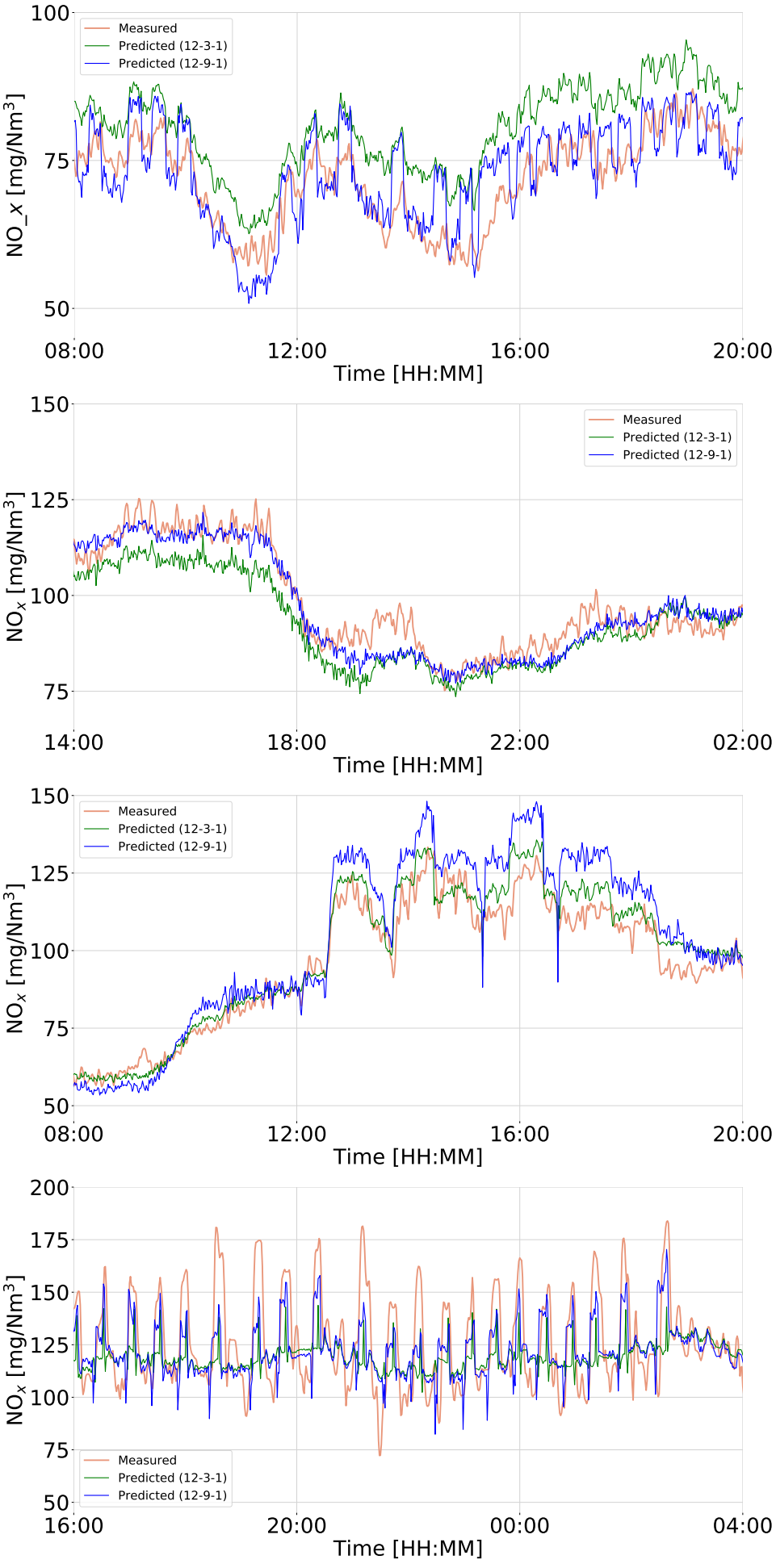


Figure 5.6. *Trends of measured and predicted NO_x emissions. Predictions are provided with the models 12-3-1 and 12-9-1.*

The first graph presents two predictions and measurement of NO_x emissions in the validation run 1. It shows that neither model can provide accurate prediction: Model with 3 HLN seems to follow variance in measured emission better than the model with 9 HLN, which overestimates the amplitude of changes. On the other hand, it is resulting in more significant biased error, around 15 mg/Nm^3, whereas the 12-9-1 has a smaller biased error. Again this biased error may be caused by the lack of measurement of fuel properties and biomass mix in model inputs.

From the next graph, validation run 2, it can be again concluded that model with 9 hidden layer neurons predicts emissions with a smaller biased error that the one with 3 HLN. However, both predictions seem to follow satisfyingly the general pattern of measured emissions.

In the third graph, validation run 3, it seems that prediction 12-3-1 follows measured emission with reasonable accuracy. Whereas model 12-9-1 seems to overestimate the amplitude of emission level changes when ammonia control is turned off, which provides peaks to prediction. All in all, it encouraging results that models can find out the dominating changes in NO_x emission trend.

When considering all four validation runs, the fourth one seems to be most interesting in terms of NO_x emissions. Because during the validation run 4, presented in the last graph, the boiler has been operated with minimum load and the nitrogen oxide emissions have been relatively high. The chart shows that measured emission levels are varying in between 100 and 175 mg/Nm^3 with quite regular intervals, which indicates that ammonia feed is turned on and of by feedforward control. It seems that both predictions can follow that dominating shape. However, when taking a closer look at the trends, it can be detected that model with 9 HLN seems to predict emissions with better accuracy.

Even though the validation scores for R^2 were slightly below 0.60, it seems that based on the trends, both models can predict the general pattern of NO_x emissions with appropriate accuracy. As well as for SO_2 model, the NO_x models seemed to provide more accurate predictions for high emission levels that the low ones. That is a good sign when considering the emission advisor application.

Modeled correlations for emissions and operational variables were studied also for NO_x models. Those were examined similarly to the SO_2 model. First emissions were predicted in an operating point occurring in input data, called base point. Then the value of the operating variable was changed while other inputs were kept constant. That was

repeated for randomly selected operation points, four of which are presented in figure 5.7 for models 12-3-1 and 12-9-1.

To get an overview of how models can predict correlations between NO_x emissions and each operational variable in different operating points, the slope of each modeled lines were examined separately. Since there is not process data available describing all predicted states, the modeled slopes were compared to the generalized correlations presented in the literature (see figure 2.9 in chapter 2). The generalised correlations are also represented in figure 5.8. ¨

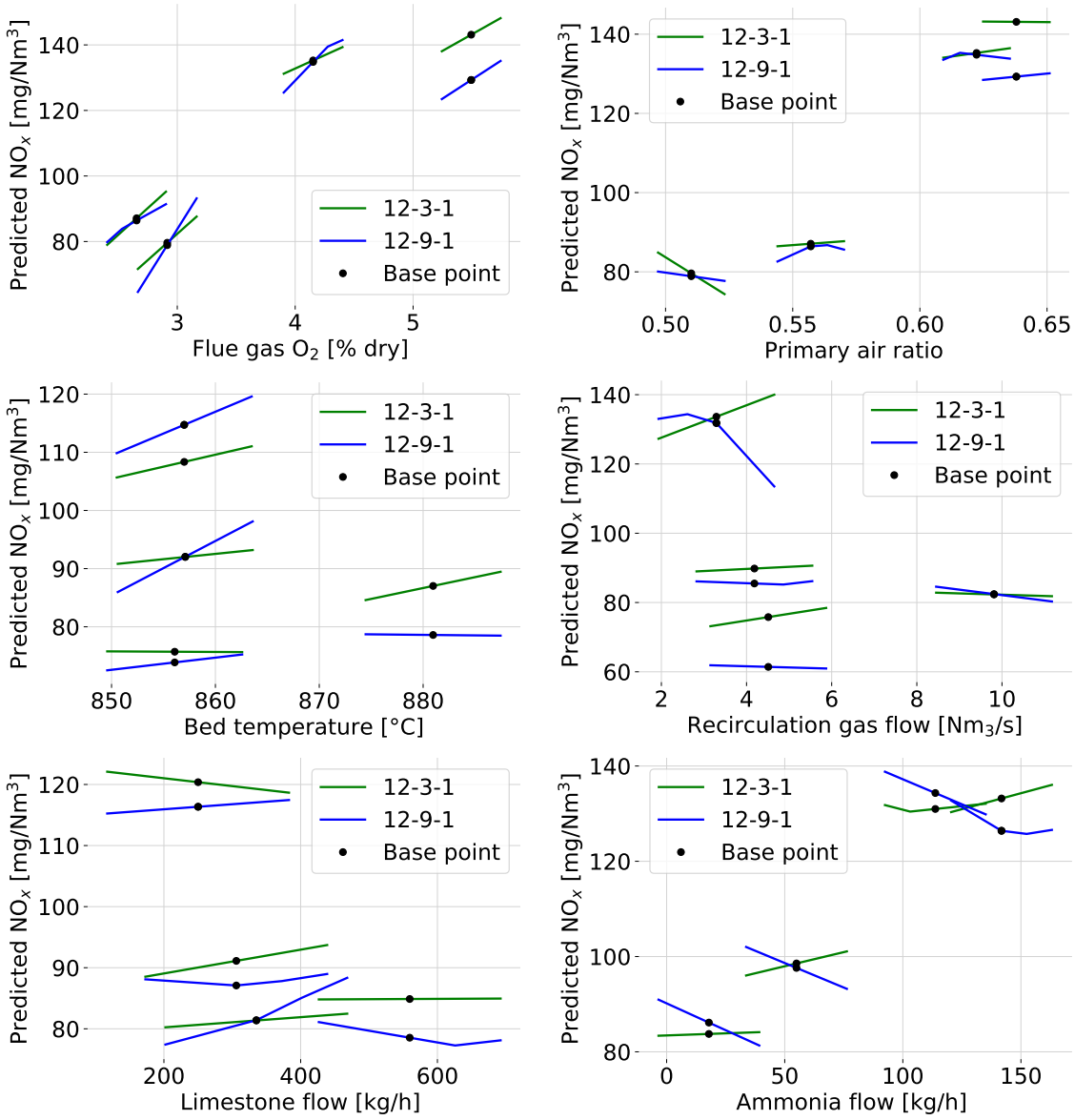


Figure 5.7. *Modeled correlations between NO_x emission and operational variables. Slope of each line shows modeled correlation around an operation point (base point).*

Parameter	Flue gas O_2	Primary air ratio	Bed temperature	Recirculation	Limestone feed	Ammonia feed
NO_x	↗	→	↗	→	↗	↘

Figure 5.8. *Generalized NOx correlations presented in the literature.*

The first graph presents modeled correlation between NO_x and flue gas O_2 content. In all four operating points, both models seem to provide consistent results: The growth in flue gas oxygen content increases NO_x emissions. This correlation is aligned with the generalized ones presented in the literature. The second graph indicates that the modeled correlation between NO_x emissions and the primary air ratio are not that unambiguous. Both models provide results which indicate that depending on the process conditions, air staging either decrease emissions or even slightly increase emissions. It is difficult to conclude which model provides results that correspond better to the ones suggested in figure 5.8.

In the third graph, the predicted correlations between NO_x and bed temperature are presented. Both models seem to predict that growth in temperature will increase nitrogen oxide emissions, which was an expected result based on figure 5.8. The four graph presents the modeled effect of flue gas recirculation. Based on the figure, it might be concluded that model 12-9-1 provides results that are better aligned with the ones presented in the literature. However, in the high emission levels, around 130 mg/Nm^3, model 12-9-1 suggests that recirculation has a surprisingly strong effect on NO_x emission reduction.

The fifth graph indicates how does the SO_2 reduction with limestone feed influence to NO_x emission level. It seems that both models provide slightly varying correlations depending on the process conditions. However, the overall outcome modeled slopes indicate that limestone addition may slightly increase NO_x emissions, as presented in the literature, too. The last graph describes the modeled correlation between ammonia feed and NO_x emissions. In this graph, the difference between models is remarkable since they both propose opposite slopes. When comparing those to the one presented in the literature, it is clear that model 12-9-1 provides better results. What is interesting that model 12-9-1 provides a result that suggests that the effect of ammonia injection may saturate when it reaches 150 kg/h.

5.1.3 CO Model

Four-fold cross-validation was also used when seeking the most suitable parameters for training and hidden layer number for CO model. These validation runs were executed with the data folds presented in figure 4.2. The best validation scores for the R^2 metric

resulted when the training parameter beta was 0.5. That value for the beta was used when examining model validation scores in different model structures.

Figure 5.9 shows how does the number of hidden layer neurons (HLN) affect validation scores for the evaluation metrics R^2 and MAE. The validation scores for R^2 seem to improve while the number of hidden layers increases. Therefore, the highest validation score for R^2 was achieved with the 10 HLN model. What is interesting in the graph is that all models, despite on 1 HLN model, provide R^2 scores higher than 0.60. This implies that models are available to predict CO emissions with unseen data with reasonable accuracy. It is surprising that based on the MAE validation scores, the model with 1 HLN seems to provide the best result, which is the opposite result of the R^2 score. A closer look at the provided predictions gave a possible explanation for this kind of model performance. In validation run 3 model with 1 HLN has converged to the solution where prediction is a constant for each operating point. Since changes in CO level are rather small, this prediction can result in small MAE error.

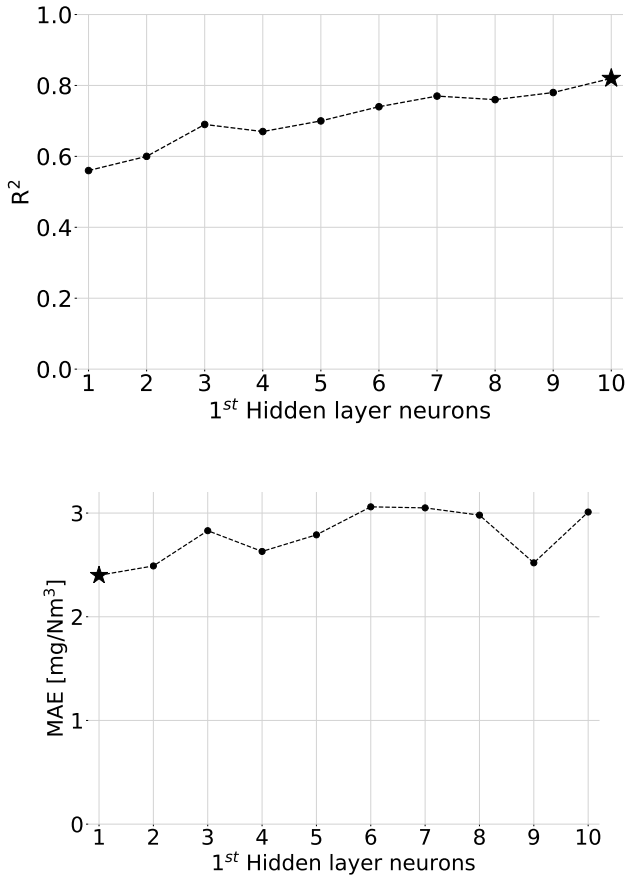


Figure 5.9. *The effect of the number of neurons on validation scores in four-fold cross validation.*

Based on the validation scores, the prediction accuracy of the two best models was examined more closely. These model structures were three HLN (8-1-1) and seven HLN (8-10-1). Unlike for the other two emission model figure 5.10 presents only one 12 hours trend of measured and predicted CO. Only one trend is shown because it was noticed that CO emission levels remain a quite similar trough each four validation. Also, CO emission level remains relatively low. It was also seen that validation scores were much more consistent for each validation run for CO model than for other two emission models.

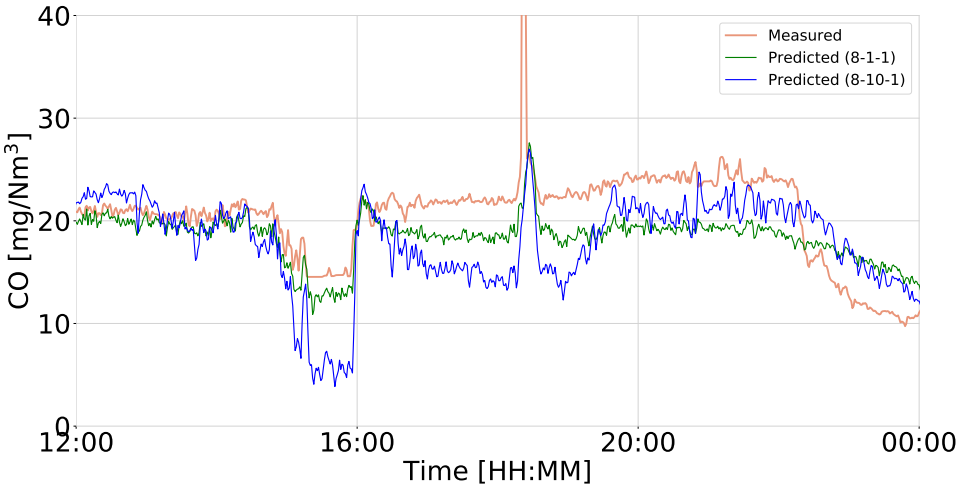


Figure 5.10. *A trend of measured and predicted CO emissions. Predictions are provided with the models 8-1-1 and 8-10-1.*

The graph presents two predictions and measurements of CO emissions in the validation run 2. It shows that at the beginning, both models follow measured emission rather well. However, at the time when there is a more substantial drop in CO emission level, both predictions seem to provide biased error: 8-1-1 smaller than 8-10-1. It appears that the model with 10 hidden layer neurons overestimates the amplitude of emission level changes, whereas 8-1-1 seems to underestimate those slightly. All in all, based on the period presented in the graph, both models appear to provide a somewhat accurate prediction.

Predicted correlations between emission and operational variables were also examined for the CO model. Based on the literature review, it was known that CO correlates most to flue gas oxygen content and bed temperature. Therefore correlations between those two and CO emissions were examined more closely as follows: First emissions were predicted in an operating point occurring in input data, called base point. Then the value of the operating variable was changed while other inputs were kept constant. That was repeated for randomly selected operation points, four of which are presented in figure

5.11 for models 8-1-1 and 8-10-1. Since there is not process data available describing all predicted states, the modeled slopes were compared to the generalized correlations presented in the literature (see figure 2.8 in chapter 2).

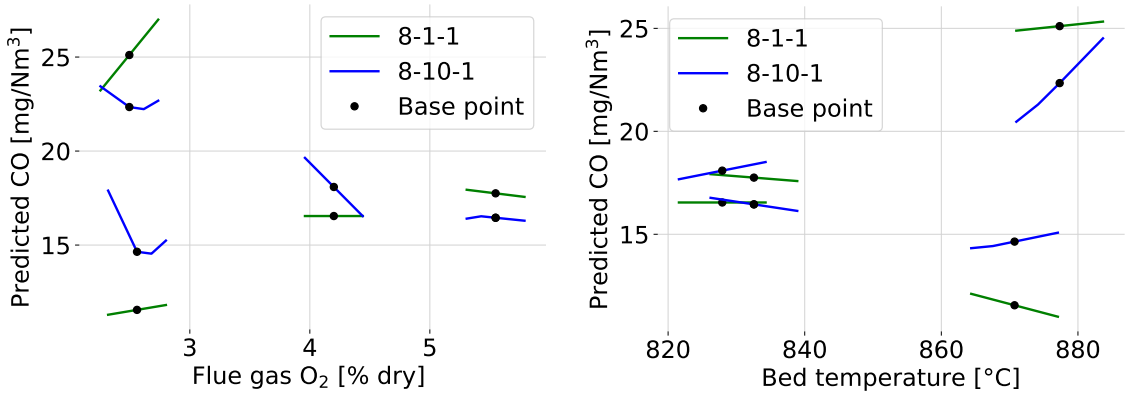


Figure 5.11. *Modeled correlations between CO emission and two operational variables. Slope of each line shows modeled correlation around an operation point (base point).*

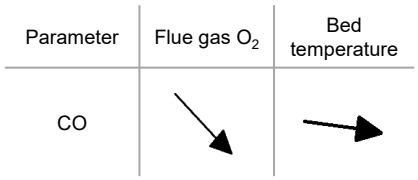


Figure 5.12. *Generalized CO correlations presented in literature.*

The first graph presents how does the flue gas oxygen content effect on CO emissions. The model with 1 HLN predicts that lowering the air excess decreases CO emissions, whereas the model with 10 HLN provides opposite results. It predicts that, especially at the low air excess levels, CO emission increases remarkably when oxygen content decreases, which correlates to the observations presented in the literature. When taking a look at the second graph, it's more challenging to draw a consistent conclusion of the predicted correlation between CO emission and bed temperature. The literature presents (figure 5.12) that CO emissions decrease when bed temperature growth.

5.1.4 Final Model Parameters

The final decision of the most suitable model structure was done based on the validation scores for R^2 and MAE metrics and modeled correlation. Also plotted trends were taken into account. The selected model and training parameters for each emission model are

summarised in the table 5.1.

Table 5.1. *The final model parameters*

Emission	Training	beta	Activation function	Model structure	R^2
SO_2	sgd-based	0.1	ReLU	8-7-1	0.56
NO_x	sgd-based	0.9	ReLU	12-9-1	0.53
CO	sgd-based	0.5	ReLU	8-10-1	0.82

The final model structure for SO_2 model was selected to be 8-7-1. The main reason to choose this architecture over the 8-3-1, which seemed to provide slightly more accurate predictions, was the fact that this model was able to give correlations that were aligned with the ones presented in the literature. For NO_x emissions, the most suitable model structure was 12-9-1 since it seemed to perform best in terms of bout prediction accuracy and correlations. Selected model architecture for CO is 8-10-1 because it provided the highest R^2 score and seemed to follow measured values rather accurately.

5.2 Overall Discussion

In this study the performance of MLP emission models were evaluated based on their prediction accuracy and their ability model correlations between emissions and operating variables that are consistent with the ones presented in the literature. Study showed that all three models were able to predict emissions with a satisfying accuracy. More importantly models provided correlations between the operational variables and emissions that were well align with the ones presented in the literature.

In this study the importance of training data quality and coverage was noted. Data used in the case study was gathered during the five months period of the reference plant operation, which ensured that the operation data represented varying process states. On the other hand, because of that the training data set included a lot of data from process states when the emissions were not high and the additives feed was off. It was also noted, that the feedback control of the additive feed likely decreased the information in data and therefore caused errors to the modeled correlations in some process states. In the future, it could be beneficial to gather data also during separate test periods in different operating points, if possible. That could increase the data quality and enable the

comparison of modeled correlations to the real data not only to the ones presented in the literature.

One of the major challenges in this study was that there was no data of fuel properties used in the model, since continuous data of those was not available. Therefore, models cannot explain the changes in emission level caused by the changes in fuel proprieties. This decreases the prediction accuracy, since emission formation, especially SO_2, is highly dependent on the fuel properties. This may at least partly explain why Krywanski2014 have presented significantly better prediction accuracy for SO_2 and NO_x MLP models, since those were trained with noncontinuous data including laboratory analysis of fuel properties. On the other hand M et al. (2011) proved that if the fuel nitrogen content remains quite stable during the biomass combustion, it is possible to predict NO_x emissions with a good accuracy by using only continuous process data.

M et al. (2011) have expressed that NO_x emissions were strongly dependent on a few input parameters and proposed data clustering to improve the model performance. The training data clustering based on the process conditions could have improved model performance in this study as well, since the model performance seemed to be more dependent on the process conditions covered in each validation fold than the model parameters. Further, the fact that models were able to predict the dominating pattern of emissions but tended to provide a biased error between the measurement and prediction implies validation set included process conditions that were not expressed in the training data. Therefore if models would be used in the emission advisor application, they should be retrained with updated process data often enough to avoid biased error.

6 CONCLUSIONS

The main objective of this work was to model the CFB boiler is emissions with a data-driven method. This study was undertaken to implement SO_2, NO_x, CO and cost model with continuous process data gathered from a reference CFB boiler and evaluate their performance. The performance was evaluated not only based on the prediction accuracy but also on how consistent the modeled correlations were compared to the ones presented in the literature.

The first research question aimed to find out what is a suitable data-driven method for modeling CFB boiler emissions. Studies suggested several different methods for boiler emission data-driven modeling. This study aims to provide a regression model and therefore, a supervised method, MLP, was selected. MLP was chosen instead of simpler regression methods, such as linear regression, because of the complexity and non-linearity of emission generation of biomass combustion. Also, dynamic neural networks have been suggested in the literature for emission modeling. However, the static method, as MLP, was chosen because it is more simple to implement and train. In the future, it could be interesting to use another approach to modeling and try other methods, for example, dynamic neural networks or unsupervised methods as well,

The second research question was set to determine which continuous process variables are needed for modeling. Based on the CFB emission theory, the suitable variables were selected from reference plant data. Process variables were chosen from fuel and air feed, additives feed, bed and chamber conditions, steam flow and flue gas. Model inputs were selected by leave one out method to find out the essential ones for model accuracy. The number of input needed were changed in between the emission models: 12 inputs for NO_x model, 8 for SO_2 and CO models. Key findings of inputs' effects on model behavior were that SO_2 was not dependent on the boiler load and CO model accuracy depended strongly on shoot blowing steam flow. Further, there was no continuous data available on fuel properties or biomass type, only fuel coal share. The lack of sufficient data of

fuel properties made it impossible to model changes caused by the variations in fuel properties. That may be an issue for further work since it is challenging to measure fuel properties in a continuous manner. In the future, it might be beneficial to utilize laboratory data besides the continuous process data in emission modeling.

The third research question was set to figure out which model parameters provide the best prediction accuracy in this case. At the beginning of models training, a suitable training algorithm and its parameters were selected. It was found out that sgd based training algorithm performed best in the models training. However, to avoid converting to a local minimum, its parameters needed to be selected carefully for each model. The main focus was to find out the best model structure, which means the number of neurons in the hidden layer(s). That was studied with four-fold cross-validation, where the model validation accuracy was evaluated with R2 score and MAE error.

NO_x model resulted in the weakest prediction accuracy since it achieved a validation score of 0.53 for the R2 score. SO_2 model provided a slightly better accuracy of 0.64 for the R2 score. CO model provided the best validation scores for R2, 0.82. These validation scores were expected since models are only able to find correlations existing in training data. Also, NO_x generation and reduction are more complex than one for the other two emissions. Measured CO were easier to predict than the other two ones since there were no significant changes in CO levels, excluding the ones caused by the shoot blowing. It was noticed that process conditions of each fold seemed to affect on model accuracy more than model structure and parameters. R2 scores varied more between the different validation folds than model structures. For example, SO_2 model provided poor R2 scores in the validation set 4, which had data from low sulfur contents, which implies that if models are used in production, those should be retrained often with data representing the latest process state, especially fuel type. Also, it could be beneficial to cluster the process state with an unsupervised method to find out whether there is new data from the process state the training data do not include.

The fourth research question was set to investigate whether the correlations between selected operating variables and modeled emissions are consistent with the ones presented in the literature. Selected operational variables were flue gas O_2, primary air ratio, bed temperature, recirculation gas flow, limestone flow and ammonia flow. It was concluded that models' ability to provide consistent results with the ones suggested in literature was dependent on the model structure. It can be concluded that for each emission model, it

was possible to find out a model structure to provide correlations corresponding to theory in most of the studied operation points. This result was encouraging since it implies then despite moderate prediction accuracy, models may be useful in predicting the correlations significant in advisor application. That provides an insight into the fifth research question, which was set to figure out whether proposed emission models are accurate enough to be used in cloud applications.

This study is a solid base point for emission advisor cloud application development. However, to achieve better prediction accuracy, models should be developed further. For example, selecting and prepossessing data whit more carefully the data quality and coverage and, though, models performance, could be improved. Still, it should be emphasized that data-driven methods are only able to model correlations existing in data. Therefore a lack of fuel quality measurement sets borders to model improvement. One approach could be to combine data-driven modeling with simple physical models to cover up the missing information. Also, it was noticed that it could be more beneficial more the application point of view to model limestone feed instead of SO_2 emissions.

REFERENCES

Abelha, P., Gulyurtlu, I. and Cabrita, I. (2008). Release Of Nitrogen Precursors From Coal And Biomass Residues in a Bubbling Fluidized Bed. *Energy & Fuels* 22.1, 363–371.

Ahrlich, S. (1975). A coal-fired fluidized bed boiler. *Fuel symp.* 1.1.

Alakangas, E. (2000). *Suomessa käytettävien polttoaineiden ominaisuuksia*. VTT. Espoo.

Alakangas, E. et al. (2018). *Growth by integrating bioeconomy and low-carbon economy*. VTT. Espoo.

Alexander, D. L. J., Tropsha, A. and Winkler, D. (2015). Beware of R2: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *J Chem Inf Model.* 55.7, 1316–1322.

Anthony, E. and Granatstein, D. (2001). Sulfation phenomena in fluidized bed combustion systems. *Progress in Energy and Combustion Science* 27.2, 215–236. ISSN: 0360-1285. URL: http://www.sciencedirect.com/science/article/pii/S0360128500000216.

Basu, P. (2006). *Combustion and Gasification in Fluized beds*. CRC Press.

Fingrid (2017). *The future of the electricity markets*. URL: https://www.fingrid.fi/en/electricity-market/the-future-of-the-electricity-markets/ (visited on 10/08/2019).

Finnish Energy (2019). *Combined heat and power generation is energy-efficient*. URL: https://energia.fi/en/energy_sector_in_finland/energy_production/combined_heat_and_power_generation (visited on 10/08/2019).

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics Data Analysis* 38.4, 367–378.

Global Market Insights (2019). *Circulating Fluidized Bed Boiler Market to hit 50 billion by 2025: Global Market Insights*. URL: https://www.globenewswire.com/news-release/2019/10/03/1924521/0/en/Circulating-Fluidized-Bed-Boiler-Market-to-hit-50-billion-by-2025-Global-Market-Insights-Inc/.

Deep Sparse Rectifier Neural Networks (2011). (14th International Conference on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, USA. Volume 15 of JMLR:WCP 15).

Golgiyaz, S., Talub, F. and Onat, C. (2019). Artificial neural network regression model to predict flue gas temperature and emissions with the spectral norm of flame image. *fuel* 255.

Gomez-Garcia, M., Pitchon, V. and Kiennemann, A. (2015). Pollution by nitrogen oxides: an approach to NOx abatement by using sorbing catalytic materials. *Environ Int* 31, 67–445.

Gungor, A. (2008). Two-dimensional biomass combustion modeling of CFB. *Fuel* 87.8, 1453–1468.

Hagan, M. and Menhaj, M. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 5.6, 989–993.

Hansen, P., Dam-Johansen, K. and Ostergaard, K. (1993). High temperature reaction between sulphur dioxide and limestone— The effect of periodically changing oxidising and reducing conditions. *Chemical Engineering Science* 48, 325–1341.

Hastie, T., Tibshirani, R. and Friedman, J. (2008). *The Elements of Statistical Learning – Data Mining, Inference, and Prediction*. 2nd ed. Springer.

Haykin, S. (2008). *Neural Networks and Learning Machines*. 3rd ed. Pearson, 44.

Hiltunen, M., Kilpinen, P., Hupa, M. and Lee, Y., eds. (1991). *11th Int. Vonf. on Fluidized Bed Combustoin*. (ASME). Vol. I0312B, 687–694.

Houshfar, E., Løvås, T. and Skreiberg, Ø. (2012). Experimental Investigation on NOx Reduction by Primary Measures in Biomass Combustion: Straw, Peat, Sewage Sludge, Forest Residues and Wood Pellets. *Energies* 5.2, 270–290.

Hupa, M. (2005). Interaction of fuels in co-firing in FBC. *Fuel* 84.10, 1312–1319. ISSN: 0016-2361. DOI: https://doi.org/10.1016/j.fuel.2004.07.018. URL: http://www.sciencedirect.com/science/article/pii/S0016236104003266.

Hupa, M., Karlström, O. and Vainio, E. (2017). Biomass combustion technology development – It is all about chemical details. *Proceedings of the Combustion Institute* 36.1, 113–134.

IEA (2017). *Technology Roadmap: Delivering Sustainable Bioenergy*. IEA. Paris. URL: http://www.iea.org/publications/freepublications/publication/Technology_Roadmap_Delivering_Sustainable_Bioenergy.pdf.

IED (2010). The reduction of national emissions of certain atmospheric pollutants (2010/75/EU) by European Union.

Ilonen, J., Kamarainen, J.-K. and Lampinen, J. (2003). Differential Evolution Training Algorithm for Feed-Forward Neural Networks. *Neural Processing Letters* 17.1, 93–105.

Koikkalainen, P., ed. (1994). *Neurolaskennan mahdollisuudet*. TEKES. Finland.

Konttinen, J., Kallio, S., Hupa, M. and Winter, F. (2013). NO formation tendency characterization for solid fuels in fluidized beds. *Fuel* 108, 238–246. ISSN: 0016-2361. DOI: https://doi.org/10.1016/j.fuel.2013.02.011. URL: http://www.sciencedirect.com/science/article/pii/S0016236113000999.

Korpela, T., Kumpulainen, P., Majanne, Y., Häyrinen, A. and Lautala, P. (2017). Indirect NOx emission monitoring in natural gas fired boilers. *Control Engineering Practice* 65, 11–25.

Koskelainen, L. and Majanne, Y. (2007). *Voimalaitosautomaatio*. Suomen Automaatioseura ry, 42–45.

Krzywański, J. and Nowak, W. (2017). Neurocomputing approach for the prediction of NOx emissions from CFBC in air-fired and oxygen-enriched atmospheres. *Journal of Power Technologies* 97.2, 75–84.

Krzywanski, J., Rajczyk, R. and Nowak, W. (2014). Model research of gas emissions from lignite and biomass co-combustion in a large scale CFB boiler. Chemical and Process Engineering. *Chemical and Process Engineering* 35.2, 217–231.

Kvalseth, T. (1985). Cautionary Note about R2. *The American Statistician* 39.4, 279–85.

Leckner, B. (1998). Fluidized bed combustion: Mixing and pollutant limitation. *Progress in Energy and Combustion Science* 24.1, 31–61.

– (2007). Co-combustion: A summary of technology. *Thermal science* 11.4, 5–40. ISSN: 0354-9836.

Leckner, B., Åmand, L., Lücke, K. and Werther, J. (2004). Gaseous emissions from co-combustion of sewage sludge and coal/wood in a fluidized bed. *Fuel* 83.4, 477–486.

Leckner, B. and Karlsson, M. (1993). Gaseous emissions from circulating fluidized bed combustion of wood. *Biomass and Bioenergy* 4.5, 379–389. ISSN: 0961-9534. DOI: https://doi.org/10.1016/0961-9534(93)90055-9. URL: http://www.sciencedirect.com/science/article/pii/0961953493900559.

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature* 7553.521, 436–444.

Lee, G. R., Gommers, R., Wasilewski, F., Wohlfahrt, K. and O'Leary, A. (2019). Scikit-learn: Machine Learning in Python. *Journal of Open Source Software* 36.4, 1237.

Liukkonen, M., Hiltunen, T., Hälikkä, E. and Hiltunen, Y. (2010). Adaptive Approaches for Emission Modeling in Circulating Fluidized Beds. *International Journal of Computer Science Emerging Technologies* 1.63.

Liukkonen, M., Hiltunen, T., Hälikkä, E. and Hiltunen, Y. (2011). Modeling of the fluidized bed combustion process and NOx emissions using self-organizing maps: An application to the diagnosis of process states. *Environmental Modelling Software* 26.5, 605–614.

Liukkonen, M. and Hiltunen, Y. (2016). Monitoring and analysis of air emissions based on condition models derived from process history. *Cogent Engineering* 3.1174182.

Lyngfelt, A., Åmand, L. and Leckner, B. (1998). Reversed air staging - a method for reduction of N20 emissions from fluidized bed combustion of coal. *Fuel* 77.9.

Lyngfelt, A. and Leckner, B. (1999). Combustion of wood-chips in circulating fluidized bed boilers — NO and CO emissions as functions of temperature and air-staging. *Fuel* 78.9, 1065–1072. ISSN: 0016-2361. DOI: https://doi.org/10.1016/S0016-2361(99)00006-X. URL: http://www.sciencedirect.com/science/article/pii/S001623619900006X.

M, L., Heikkinen, M., Hiltunen, T., Hälikkä, E., Kuivalainen, R. and Hiltunen, Y. (2011). Artificial neural networks for analysis of process states in fluidized bed combustion. *Energy* 36.1, 339–347.

MacGregor, J. and Cinar, A. (2012). Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven methods. *Computers and Chemical Engineering* 47, 111–120.

Teollinen internet uudistaa palveluliiketoimintaa ja kunnossapitoa (2018). Vol. 102. 2, 1645–1656.

Mathieu, J., Tzanis, L., Soulard, M., Patarin, J., Vierling, M. and M., M. (2013). Adsorption of SOx by oxide materials: A review. *Fuel Processing Technology* 114, 81–100.

Ministry of the Environment (2019). National Air Pollution Control Programme 2030. 2019:14, 96. URL: http://urn.fi/URN:ISBN:978-952-361-020-0.

Moller, M. (1993). Efficient Training of Feed-Forward Neural Networks. *DAIMI Report Series* 22.

Nussbaumer, T. (2003). Combustion and co-combustion of biomass: fundamentals, technologies, and primary measures for emission reduction. *Energy and Fuels* 17.6, 1510–1521.

Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.

Qian, F., Chyang, C., Chiou, J. and Tso, J. (2011). Effect of Flue Gas Recirculation (FGR) on NOx Emission in a Pilot-Scale Vortexing Fluidized-Bed Combustor. *Energy and Fuels* 25, 5639–5646.

Raiko, R., Saastamoinen, J., Hup, a. M. and Kurki-Suonio, I. (2002). *'Potto ja palaminen'*. Teknistieteelliset akatemiat, 300–393.

Rayaprolu, K. (2009). *Boilers for Power and Process, 1.st edition*. Boca Raton, 745 p.

Skalska (2010). Trends in NOx abatement: A review. *Science of The Total Environment* 408.19, 3976–3989.

Solomatine, D., See, L. and Abrahart, R. (2008). Data-Driven Modelling: Concepts, Approaches and Experiences. *Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications*. Ed. by R. J. Abrahart, L. M. See and D. P. Solomatine. Berlin, Heidelberg: Springer Berlin Heidelberg, 17–30.

Spliethoff, H. (2010). *Power Generation from Solid Fuels*. Springer-Verlag Berlin Heidelberg, 234–306.

Tarelho, L., Matos, M. and Pereira, F. (2005). The influence of operational parameters on SO2 removal by limestone during fluidised bed coal combustion. *Fuel Processing Technology* 86.12, 1385–1401. ISSN: 0378-3820. DOI: https://doi.org/10.1016/j.fuproc.2005.03.002. URL: http://www.sciencedirect.com/science/article/pii/S0378382005000512.

Transparency Market Research (2014). *Transparency Market Research Circulating Fluidized Bed (CFB) Boilers Market - Global Industry Analysis, Size, Share, Growth, Trends, and Forecast 2015 - 2023*. URL: https://www.transparencymarketresearch.com/cfb-mar-ket/.

Valmet (2019). *NOx reduction*. URL: https://www.valmet.com/energyproduction/air-emission-control/nox-reduction/ (visited on 11/12/2019).

Vermeulen, I., Block, C. and Vandecasteele, C. (2012). Estimation of fuel-nitrogen oxide emissions from the element composition of the solid or waste fuel. *Fuel* 94, 75–80. ISSN: 0016-2361. DOI: https://doi.org/10.1016/j.fuel.2011.11.071. URL: http://www.sciencedirect.com/science/article/pii/S001623611100785X.

Wang, Z., Zhou, J., Zhu, Y., Wen, Z., Liu, J. and Cen, K. (2007). Simultaneous removal of NOx, SO2 and Hg in nitrogen flow in a narrow reactor by ozone injection: Experimental results. *Process Technol* 88, 23–817.

Wu, Y.-c. and Feng, J.-w. (2018). Development and Application of Artificial Neural Network. *Wireless Personal Communications* 102.2, 1645–1656.