Applied Energy 254 (2019) 113668

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Carbon monoxide emission models for small-scale biomass combustion of wooden pellets

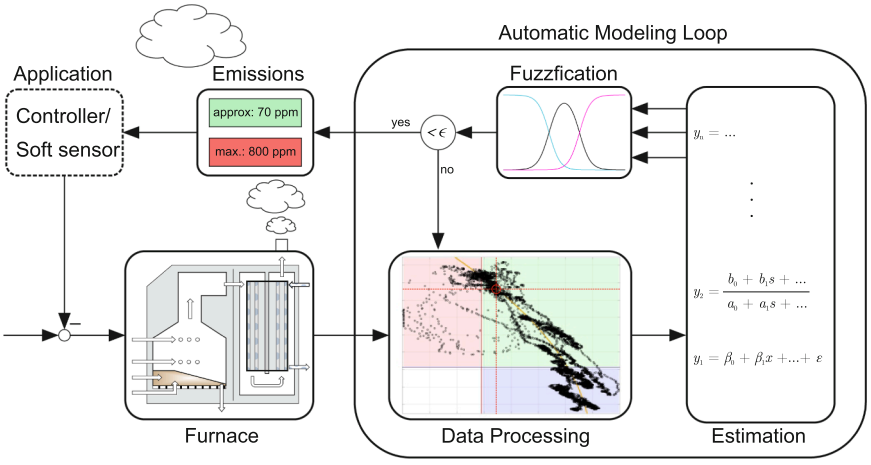
Lukas Böhler[a,*], Gregor Görtler[b], Jürgen Krail[b], Martin Kozek[a]

[a] Vienna University of Technology, Institute of Mechanics and Mechatronics, 1060 Vienna, Austria
[b] Fachhochschule Burgenland GmbH, 7423 Pinkafeld, Austria

HIGHLIGHTS

- Nonlinear kinetic model for the estimation of carbon monoxide in biomass combustion.
- Automatic partitioning and estimation concept for carbon monoxide emissions.
- Fuzzy framework for the combination of different black-box structures.
- Resulting model structure is applicable for controller and carbon monoxide soft sensors.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Biomass
Carbon monoxide
Emission
Estimation
Modeling

ABSTRACT

Tighter legal emission limits require means to prevent releasing harmful substances into the atmosphere during the combustion of biomass. Economic considerations suggest to meet these restrictions by improving the ability to predict and therefore prevent emissions, which can be done by improved control algorithms. This work presents different methods to obtain models for the prediction of carbon monoxide emissions in a small-scale biomass combustion furnace for wooden pellets. The presented models are intended for an application in model based control, either as part of the underlying model or for carbon monoxide soft sensing and fault detection. The main focus is on simple structures which can be handled by the already existing hardware of the furnaces. Different black-box models and a kinetic process model are introduced and compared. The black-box models are based on the measured flue gas oxygen concentration and the combustion temperature, since these measurements are typically available even for smaller plants. The obtained models are validated with measured data in order to find the most suitable structures, of which combined fuzzy black-box models show the most promising results. The presented methodology can be readily applied to the investigated furnace. However, the model parameters have to be adapted for other plants.

* Corresponding author.
E-mail addresses: lukas.boehler@tuwien.ac.at (L. Böhler), Gregor.Goertler@fh-burgenland.at (G. Görtler), Juergen.Krail@fh-burgenland.at (J. Krail), martin.kozek@tuwien.ac.at (M. Kozek).

https://doi.org/10.1016/j.apenergy.2019.113668
Received 28 March 2019; Received in revised form 25 July 2019; Accepted 1 August 2019

1. Introduction

The combustion of wooden biomass is a renewable and carbon dioxide (CO_2) neutral source of energy. Alongside the formation of CO_2 other harmful gaseous emissions like carbon monoxide (CO), nitrogen oxides (NO_x) and particulate matter are released into the atmosphere. Legal restrictions and an environmental conscious combustion require to minimize these emissions. Secondary measures for flue gas cleaning can be expensive additions especially to small-scale furnaces, whereas improving the existing combustion control can be more cost efficient. This however requires capable and advanced control strategies, as well as means to anticipate the formation of pollutants during the process. This work therefore aims to provide and compare different data driven model approaches and a kinetic model tailored for the application in model based control for small-scale biomass combustion furnaces with limited computational capacities. The provided models are intended to be utilized either for soft-sensing and subsequent CO sensor fault detection or as part of the underlying model for the controller. A modeling framework is presented that is able to cover the considered range of operation and dynamic load changes of the investigated furnace for the combustion of wooden pellets.

Capturing the different thermo-chemical processes occurring during the combustion of biomass is a difficult task and first principle models often have to be designed anew for every specific furnace type [1]. Stoichiometric white-box descriptions behind the combustion of wooden biomass, including the formation of carbon monoxide and other gaseous emissions are described in literature. An overview is given for example in the handbook of combustion [2]. Stoichiometric approaches help to understand the physical phenomena behind the combustion process, but their application requires permanent knowledge of the fuel composition and the fuel flow entering the system [3], which is rarely the case for small-scale furnaces. The same restrictions apply to kinetic or non-stoichiometric approaches [4], if detailed mass and energy balances are necessary. Models based on partial differential equations as in [5], even if they are compactly formulated as in [6], can yield satisfying results, but are often too complex for a direct application in model based control [7]. A new kinetic model approach to estimate the CO emissions is therefore presented.

Black-box models on the other hand are based on measured input and output data without default specifications for their internal structure. Hence, any available measurement or signal can directly be used as input to describe a specific output. Measuring the CO concentration in the flue gas reliably can be very expensive for small-scale furnaces and is currently not standard, however it is required to meet the legal restrictions. If no sensor is available, an emission estimation model can be used to predict potential violations and to increase the performance of the plants. In biomass combustion, the classification of measured input–output data has resulted in characteristic CO concentration diagrams as presented in [8] for different furnaces and in [9] for different fuels. These diagrams display the carbon monoxide levels over the excess air ratio λ and/or the combustion temperature. In [10] CO-λ diagrams are used to describe the effects to the CO formation when the fuel composition changes. A rule based fuzzy control concept built on these characteristics has been derived in [11] for example. Another rule based control strategy for CO emission reduction is presented in [12]. These approaches do not rely on explicit emission models, but rather directly formulate the control rules depending on the current CO-λ combinations. The prediction of the CO concentration due to transient processes however is not displayed by these diagrams. Black-box models can be utilized to close the gap between these observations and the requirements for advanced control algorithms.

Data driven models for the estimation of CO however have been sparsely represented in literature compared to the extensive research conducted for the estimation of NO_x. In [1] for example, neuro-fuzzy models with a distributed logic processor structure are introduced to estimate the NO_x concentration of a fluidized bed combustion plant. The prediction of multiple emissions with regression based neural networks is presented in [13] for a 600 MW utility boiler. Machine learning methods have been applied as well, like deep learning techniques to model NO_x based on flame radical image processing [14]. Exploratory surveys of different black-box structures again for the estimation of nitrogen oxides are presented in [15] and in [16]. While [15] derived models for the NO_x estimation in natural gas fired boilers and [16] for a coal-based boiler, both provide prediction models for controller or soft-sensor applications. In [17] a NARX-neural network is used to estimate CO and NO_x concentrations of a medium-scale coal furnace. Although most of these approaches focus on nitrogen oxides, they can act as a guideline for the deployment of data driven CO models.

This work is concerned with different approaches for the estimation of carbon monoxide emissions in a small-scale biomass combustion furnace for wooden pellets. The derived models are intended for the application as a CO soft sensor or as part of a model predictive controller (MPC). Due to the considered furnace size and the intended controller integration, a major requirement for the obtained models is that they are computationally manageable. The introduced static and dynamic structures are based on the insight gained from literature and the available measured data of the investigated furnace. The main result is a self-tuning fuzzy framework which can consist of a composition of local black-box models with different structures, including regression, transfer function or Hammerstein approaches. A neural network serves as a benchmark, since its performance has been shown in various publications [18]. For all of the local fuzzy models a global counterpart is trained to highlight the benefits of splitting the partitioning space. The presented analytical model is based on kinetic Arrhenius type equations with a special focus on the reactions of carbon and oxygen. This provides a tailored approach for the estimation of CO which is still applicable to the limited available hardware. The results are validated and compared with measured data.

This paper consists of the following sections: First, a short introduction of the process is given that continues with the analytical process model. Then, the black-box structures and fuzzy methods are introduced. Next, the observations from the measured data and the inferred parametrization of the models is presented. Finally, the models are validated and the results are discussed.

2. Analytical approach

This section presents the basic furnace structure and an analytical approach to estimate the carbon monoxide concentration of the flue gas. Although data driven approaches are getting more attention recently, kinetic models are capable to meet the requirements for model based control as well. In [5] such an analytical approach is presented, which is derived for the estimation of NO_x in a coal-fired boiler. The model presented in this section however is even simpler and not based on distributed process descriptions. The nomenclature of this section is given in Table 1.

2.1. Furnace structure

The considered plant is a small-scale biomass grate furnace with two staged combustion. A schematic of the plant is given in Fig. 1. In the first stage on the grate, the fuel and the primary air react in a sub-stoichiometric combustion. The intermediate gaseous products and volatiles are fully oxidized in the second stage in the freeboard above the grate. Small-scale furnaces (up to 1 MW) differ in their designs and therefore their available measurements. The application of oxygen sensors in the flue gas stream and temperature measurements in the freeboard however are standard for automated furnaces. They provide essential process feedback and are therefore important parameters for control strategies considering emission reduction [19]. Controllable inputs are the intake air fan, a secondary air valve for a more precise

Table 1
Nomenclature of the analytical equations for this section.

Symbol	Description	Unit
A	frequency factor	1/s
c	concentration of species C,O,H,N	1
E	activation energy	J/mol
$\dot{m}$	mass flow	kg/s
n	molar quantity	mol
R_g	universal gas constant	J/(mol K)
T	temperature	K
V	volume	m^3
$\dot{V}$	volume flow	m^3/s
x, y, z	molecular ratios	1

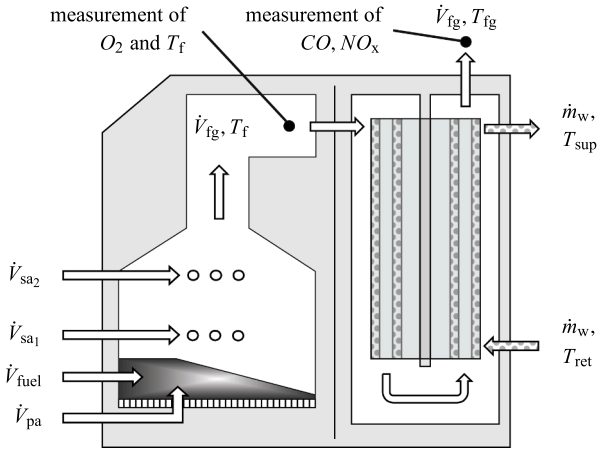


Fig. 1. Basic structure of the investigated 100 kW grate combustion furnace with the location of the measurements of flue gas oxygen O_2, freeboard temperature T_f, carbon monoxide CO and nitrogen oxide NO_x concentrations. Volume flows: primary air $\dot{V}_{pa}$, secondary air 1 and 2 $\dot{V}_{sa1,2}$, biomass fuel $\dot{V}_{fuel}$, flue gas $\dot{V}_{fg}$ and water mass flow $\dot{m}_w$ of the heating circuit. Temperatures: flue gas T_{fg}, supply water T_{sup} and return water T_{ret}.

burnout control in the freeboard and the fuel conveyor speed. With those inputs, the air and fuel volume inflows are calculated which are necessary for the derivation of the analytical model presented in this section. The mass flow of the water in the heating circuit is commonly not available, in contrast to the return and supply temperatures, but is also not required for the models in this work. Solely the flue gas volume flow is often not measured, but can be estimated from the fuel stoichiometry and the oxygen measurement, e.g. as in [20]. An alternative approach to estimate the flue gas volume flow in steady state is to simply sum up the inputs

$$\dot{V}_{fg} = \dot{V}_{fuel} + \dot{V}_{pa} + \dot{V}_{sa} + \dot{V}_{a,leakage} \quad (1)$$

or to use the suggested experimental relation

$$\dot{V}_{fg} = (\dot{V}_{pa} + \dot{V}_{sa})^{k_{fg}} \quad (2)$$

for more dynamic load changes, with $\dot{V}_{sa}$ as the summarized secondary air flow, $\dot{V}_{a,leakage}$ as the air sucked into the furnace due to leakages and k_{fg} as a furnace specific factor, which is determined by measurements assuming that the stoichiometric relation between $\dot{V}_{fuel}$ and $\dot{V}_{pa} + \dot{V}_{sa1}$ is approximately constant. The remaining indexes are as described in Fig. 1. Eqs. (1) and (2) however are only correct if the density of the air and the flue gas are constant within small limits, which is typically the case for the investigated furnace.

2.2. Considered reactions

The combustion of biomass consists of a series of complex reactions as presented in [2]. To derive a compact kinetic model, simplifying some of these thermo-chemical relations is necessary. For wooden pellets according to the EN Plus standard it is assumed that the water mass fraction of the wet fuel and the elemental composition of the dry fuel remain constant. Additionally, the amount of trace elements like sulfur or potassium is neglected. For the kinetic approach only elemental reactions of carbon (C), oxygen (O), hydrogen (H) and nitrogen (N) are considered. The relevant reactions on the grate are.

R1: $2\,C + O_2 \rightarrow 2CO$

R2: $2xC + yH_2 \rightarrow 2C_xH_y$

R3: $N_2 + xO_2 \rightarrow 2NO_x$

resulting in CO, hydrocarbons (C_xH_y) and NO_x given the substoichiometric air to fuel ratio $\lambda_{grate} < 1$. A more detailed separation into pyrolysis and char combustion on the grate as in [6] or [21] is not considered in this model. The formation of NO_x is divided according to [22] into three different mechanisms for biomass combustion: thermal, prompt and fuel NO_x. The formation of thermal NO_x is not considered, because the necessary temperature of the hot gas mixture in the freeboard is not high enough (<1000 K) for this reaction. Prompt NO_x formation takes place, if hydrocarbons are present in fuel-rich regions, but doesn't need to be considered for biomass if the combustion temperatures are beneath 1100 K either. The major part of the measured NO_x concentration in the flue gas therefore results from the fuel nitrogen. For the given furnace, the flue gas NO_x can be described by a linear relation with the amount of fuel fed into the furnace. The reaction for water steam (H_2O)

R4: $2H_2 + O_2 \rightarrow 2H_2O$

can take place either on the grate or in the freeboard. The remaining reactions of the intermediate combustion products in the freeboard are

R5: $2CO + O_2 \rightarrow 2CO_2$

R6: $C_xH_y + zO_2 \rightarrow xCO_2 + y/2H_2O$

with $z = x + y/4$. Assuming a complete burnout due to sufficient oxygen ($\lambda > 1$), residence time, temperature and turbulence, the carbon monoxide molecules in R5 and the hydrocarbons in R6 are fully oxidized. The CO concentration in the flue gas however increases rapidly if the furnace is not operated under ideal conditions and therefore indicates an incomplete combustion.

2.3. Modeling of the combustion

Based on equations R1 to R6, a kinetic grey-box model is derived for the estimation of CO. The model consists of individual ordinary first order differential equations for each involved species. Eventually only the carbon monoxide concentration is of interest, although the measurable flue gas oxygen and carbon dioxide concentrations are outputs of the model as well. The nitrogen of the fuel from R3 can be eliminated from the set of equations if the reacting oxygen is accounted for. Assuming that mostly nitrogen oxide (NO) is generated, the reduced amount of available O_2 molecules $O_{2,red,NO}$ for the formation of steam and CO_2 can be expressed as

$$O_{2,red,NO} = \frac{n_{N_2}}{n_C + 0.5n_{H_2} + n_{N_2}}.$$

Similarly, the hydrogen is assumed to completely react to water, which needs a significant amount of oxygen. The reduced amount of oxygen O_{2,red,H_2O} can be accounted for as

$$O_{2,red,H_2O} = \frac{0.5n_{H_2}}{n_C + 0.5n_{H_2} + n_{N_2}}. \quad (3)$$

The remaining O_2 molecules are then available for the formation of CO and CO_2. This has to be considered in the combustion air feed. The volume flows $\dot{V}_{pa}$ of primary and $\dot{V}_{sa}$ of secondary air are then

$$\dot{V}_{pa} = 0.21\dot{V}_{pa,in}\left[1 - O_{2,red,NO} - O_{2,red,H_2O}k_{shift}\right]$$

and

$$\dot{V}_{sa} = 0.21\dot{V}_{sa,in}\left[1 - O_{2,red,H_2O}(1 - k_{shift})\right]$$

where $\dot{V}_{pa,in}$ and $\dot{V}_{sa,in}$ are the measured or calculated primary and secondary air volume flows into the furnace and k_{shift} is a experimentally determined factor between 0 and 1 that shifts the oxygen consumption due to the formation of H_2O from R4 and R6 between grate and freeboard. The concentration of oxygen on the grate $c_{O_{2,g}}$ is

$$\frac{dc_{O_{2,g}}}{dt}V_1 = \dot{V}_{fuel,O_2} + \dot{V}_{pa} - c_{O_{2,g}}\dot{V}_0 - R_{1,kin} \tag{4}$$

with V_1 as the virtual grate volume, $\dot{V}_{fuel,O_2}$ the volume flow of oxygen contained in the fuel, $\dot{V}_0$ the volume flow from the grate into the freeboard and $R_{1,kin}$ as the kinetic reaction rate. The volume flow $\dot{V}_0$ is given as

$$\dot{V}_0 = \dot{V}_{fg} - \dot{V}_{sa}$$

and the kinetic reaction rate based on R1 as

$$R_{1,kin} = A_1 V_1 c_{O_{2,g}} c_C^2 \exp\left(\frac{-E_1}{R_g T}\right)$$

with the experimentally determined frequency factor A_1 and activation Energy E_1. The carbon part is given by

$$\frac{dc_C}{dt}V_1 = \dot{V}_{fuel,C} - c_C\dot{V}_0 - R_{1,kin} \tag{5}$$

with $\dot{V}_{fuel,C}$ as the volume flow of carbon into the system. The concentration of carbon monoxide from the grate c_{CO_g} is obtained as

$$\frac{dc_{CO_g}}{dt}V_1 = 2R_{1,kin} - c_{CO_g}\dot{V}_0. \tag{6}$$

Due to the simplified reaction kinetics an additional state is introduced in the freeboard to account for the dynamic behavior of the process. This state represents an all-pass denoted as R_a and is defined as

$$\frac{dR_a}{dt}t_a = R_{2,kin} - R_a$$

with t_a as an experimentally determined time constant and the kinetic reaction rate $R_{2,kin}$ based on R5 as

$$R_{2,kin} = A_2 V_2 c_{O_{2,f}} c_{CO_f}^2 \exp\left(\frac{-E_2}{R_g T}\right)$$

with the experimentally determined frequency factor A_2 and activation Energy E_2. The oxygen concentration in the freeboard $c_{O_{2,f}}$ is

$$\frac{dc_{O_{2,f}}}{dt}V_2 = c_{O_{2,g}}\dot{V}_0 + \dot{V}_{sa} - c_{O_{2,f}}\dot{V}_{fg} - (1 - k_a)R_{2,kin} - k_a R_a \tag{7}$$

with the volume V_2 and k_a as the experimentally determined all-pass gain for the system. The concentration of carbon dioxide $c_{CO_{2,f}}$ in the freeboard is obtained from

$$\frac{dc_{CO_{2,f}}}{dt}V_2 = 2R_{2,kin} - c_{CO_{2,f}}\dot{V}_{fg} - k_a(R_{2,kin} - R_a) \tag{8}$$

and the concentration of carbon monoxide c_{CO_f} is

$$\frac{dc_{CO_f}}{dt}V_2 = c_{CO_g}\dot{V}_0 - c_{CO_f}\dot{V}_{fg} - R_{2,kin}. \tag{9}$$

The carbon monoxide in the flue gas originates from R1, but can also be a product of the incomplete oxidation of hydrocarbons at any stage. Modifying R6 to account for CO leads to

R7: $\quad C_xH_y + zO_2 = \alpha(xCO_2 + y/2H_2O) + x(1 - \alpha)CO$

with the virtual stoichiometric factor

$$\alpha(T_f, O_2) = \{\alpha \in \mathbb{R} | 0 \leqslant \alpha \leqslant 1\}$$

which shifts the reaction from CO_2 to CO depending on the measured temperature and oxygen concentration. For a complete burnout α is 1. Since no additional oxygen is required for H_2O due to Eq. (3) and z depending on α now as well, R7 is reformulated to

R8: $\quad C_x + z_\alpha O_2 = x(\alpha CO_2 + (1 - \alpha)CO)$

with $z_\alpha = 1/(2 - \alpha)$. This is the basis for the kinetic reaction describing the hydrocarbon decomposition

$$R_{3,kin} = A_3 V_2 c_C c_{O_{2,f}}^{z_\alpha} \exp\left(\frac{-E_3}{R_g T}\right)$$

with the experimentally determined frequency factor A_3 and activation Energy E_3. The concentration of O_2, CO_2 and CO measured in the flue gas is then given by:

$$\frac{dc_{O_2}}{dt}V_2 = (c_{O_{2,f}} - c_{O_2})\dot{V}_{fg} - R_{3,kin} \tag{10}$$

$$\frac{dc_{CO_2}}{dt}V_2 = (c_{CO_{2,f}} - c_{CO_2})\dot{V}_{fg} + \alpha R_{3,kin} \tag{11}$$

$$\frac{dc_{CO}}{dt}V_2 = (c_{CO_f} - c_{CO})\dot{V}_{fg} + (1 - \alpha)R_{3,kin} \tag{12}$$

These equations are basically extending Eqs. (7)–(9) with the additional reaction for the hydrocarbons. The kinetic model now consists of 10 differential and a varying number of algebraic equations, depending on whether the measurements of the fuel and combustion air are available or not. In any case, Eqs. (4)–(12) represent the introduced analytical model. Some assumptions made in this section are based on using wooden pellets as fuel, which is the limiting factor for the applicability of this model.

3. Data driven models

In this section the methods and the structures for the data driven models are presented. The black-box models are based on the flue gas oxygen concentration and the freeboard temperature, as suggested in [19]. Other inputs have been considered as well, but have been discarded after a cross-correlation analysis. This work focuses on a fuzzy approach based on different local models (submodels), which means that not the entire partitioning space is covered by one model alone. To emphasize the fuzzy approach, additional global models are parametrized for comparison. A global model is therefore a single model that covers the entire partitioning space on its own.

3.1. Overview

The considered black-box structures are a multivariate static regression (SR) model, a dynamic ARX model, a feed forward neural network and a Hammerstein model (HM). Additionally, a transfer function model which uses the SR as input is introduced. This structure is in fact a Hammerstein model as well, but with a different description of the input nonlinearity. The regression models are linear in the parameters, but can be nonlinear in the inputs. The parameters for the SR and the ARX are therefore estimated with least squares techniques. Linear regression and ARX models have been applied in comparable investigations for the estimation of NO_x emissions in [15,16] respectively. The neural nets are multi-layer perceptron (MLP) networks with one hidden layer and one output layer as suggested in [15] for the estimation of NO_x and trained with the Bayesian regularization method [23]. Neural networks are powerful structures and have been applied various times in literature to model combustion processes [18] and have to be considered, if different data driven models are compared. Hammerstein approaches are not very common for emission estimation, which is why the applicability of these structures is investigated in this work as well.

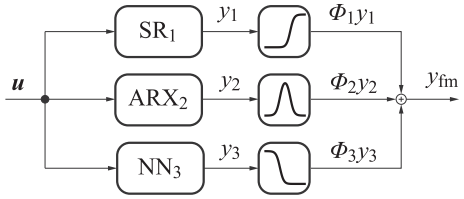


Fig. 2. Block representation of a potential fuzzy model with mixed static and dynamic submodels of different structure. The index fm for the output y indicates that y_{fm} is the result of a fuzzy model. The weighting or validity functions Φ determine how much a local model contributes to the output.

3.2. Fuzzy modeling

By splitting the data into local zones, different model structures and types can be applied, e.g. mixing dynamic and static or linear and nonlinear submodels. The partitioning space is determined by the oxygen concentration $u_1 = O_2$ and the freeboard temperature $u_2 = T_f$. An explanatory block diagram of how a fuzzy model can look like is given in Fig. 2, where a SR, ARX and a NN model are mixed. The CO concentration is therefore a weighted sum of the outputs of the local models that depends on the inputs u_p with p = 1,2 and is obtained based on [24] as

$$y_{fm} = \sum_{m=1}^{3} \Phi_m(u_p) y_m \tag{13}$$

where $\Phi_m(u_p)$ are the validity functions and y_m are the outputs of the m submodels with m = 1,2,3. The selected number of submodels will be discussed in Section 4. The validity functions Φ_m are defined as

$$\Phi_m(u_p) = \frac{\mu_m(u_p)}{\sum_{q=1}^{3} \mu_q(u_p)} \tag{14}$$

where $\mu_m(u_p) = \mu_q(u_p)$ are the Gaussian membership functions μ defined as

$$\mu(u_p) = \exp\left(-\frac{1}{2(1-r_m^2)}\left(\frac{(u_1 - cp_{m,1})^2}{\sigma_{m,1}^2} + \frac{(u_2 - cp_{m,2})^2}{\sigma_{m,2}^2} - \frac{2r_m(u_1 - cp_{m,1})(u_2 - cp_{m,2})}{\sigma_{m,1}\sigma_{m,2}}\right)\right) \tag{15}$$

where $cp_{m,p}$ are the centerpoints, $\sigma_{m,p}$ the spreads and r_m are the correlation coefficients. The correlation coefficient rotates the local membership function in the input space and is defined between -1 and 1. The locations of the centerpoints, the spreads and the correlation coefficients are degrees of freedom for the parametrization of the fuzzy model.

3.3. Model structures

The following structures are defined for the global and local models alike, though for the global structures the index m can be omitted. The multivariate static regression models are formulated as

$$y_{m,SR}(k) = \theta_0 + \theta_1 x_1(k-1) + \theta_2 x_2(k-1) + \ldots + \theta_j x_j(k-1) + \epsilon(k) \tag{16}$$

where $y_{m,SR}(k)$ is the output of a model m at sampling time k and θ_j are the j parameters to be estimated. The j regressors x_j have a fixed time delay of one sample, which is also applied to the inputs for the neural network. The x_j can be any combination of the inputs $u_1(k-1)$, $u_2(k-1)$ or one of the two with exponents greater than 1. The first regressor $x_0(k-1)$ is always 1. The additive process noise $\epsilon(k)$ is a stochastic variable assumed to have zero mean and variance $\sigma_\epsilon^2 = E\{\epsilon^2(k)\}$. In compact form, Eq. (16) can be written as

$$\boldsymbol{y}_{m,SR} = \boldsymbol{X}_{m,SR}\boldsymbol{\theta}_{m,SR} + \boldsymbol{\epsilon}$$

where $\boldsymbol{X}_{m,SR}$ is the $N\times$ j matrix of regressors with N as the considered number of input–output data, $\boldsymbol{\theta}_{m,SR}$ the concatenated j $\times$ 1 vector of parameters and $\boldsymbol{\epsilon}$ the $N \times 1$ noise vector. The distinction into local and global regression models is illustrated in Fig. 3. A typical formulation of the ARX model [25] is

$$\begin{aligned} y_{m,ARX}(k) &= a_{m,1}y_m(k-1) + \ldots + a_{m,n_a}y_m(k-n_a) + \\ &+ b_{m,1,p}u_p(k-n_{k,p}) + \ldots + \\ &+ b_{m,n_{b,p},p}u_p(k-n_{k,p}-n_{b,p}+1) \end{aligned} \tag{17}$$

where n_a is the number of considered past outputs and $n_{b,p}$ is the number of parameters for each input with the according time delay $n_{k,p}$. Again in compact form, Eq. (17) can be rewritten as

$$\boldsymbol{y}_{m,ARX} = \boldsymbol{X}_{m,ARX}\boldsymbol{\theta}_{m,ARX} + \boldsymbol{\xi}$$

with the filtered additive process noise

$$\boldsymbol{\xi} = \frac{1}{A(z^{-1})}\boldsymbol{\epsilon}$$

where z^{-1} is the backward shift operator and A the denominator polynomial of the ARX model. The solution for the parameters in $\boldsymbol{\theta}$ for the SR and the ARX can be obtained by solving [26]

$$\boldsymbol{\theta} = \left(\sum_{i=1}^{n} \boldsymbol{X}_i^T \boldsymbol{X}_i\right)^{-1} \sum_{i=1}^{n} \boldsymbol{X}_i^T \boldsymbol{y}_i \tag{18}$$

where i is the i-th separate data segment of the n considered measurements for the parameter estimation of each of the m zones. The Hammerstein structure consists of a memoryless nonlinear input block and a dynamic linear transfer function block. The static input nonlinearity is described by a piecewise linear function. The configuration in block form is presented in Fig. 4.

The output $y_m(k)$ of a Hammerstein model is obtained as

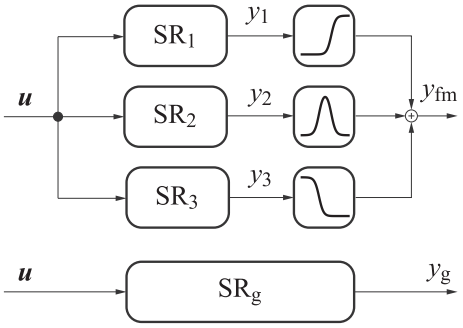


Fig. 3. Block representation of a fuzzy model consisting only of static regression models (top) compared to the global SR model (bottom). This block representation also applies to the ARX and NN models. The index g indicates the output of the global model.

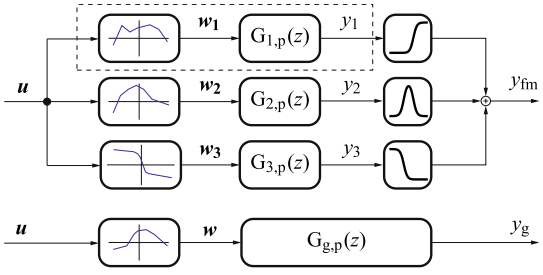


Fig. 4. Block representation of the Hammerstein models with a static nonlinear input block, the inner inputs $\boldsymbol{w}_m$ and a linear transfer function block consisting of two transfer functions $G_{m,p}$ for each of the inputs p = 1,2. The dashed rectangle highlights the general structure of a Hammerstein model.

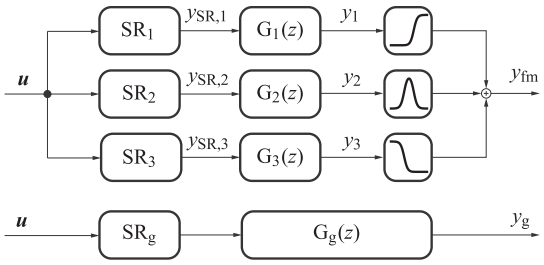


Fig. 5. Block representation of the regression and transfer function model where the input of the additional transfer function is a SR model.

$$y_{\mathrm{m,HM}}(k) = G_{\mathrm{m},1}(z)w_1(k) + G_{\mathrm{m},2}(z)w_2(k)$$

with the inner inputs $w_1(k)$ and $w_2(k)$ and the discrete transfer functions $G_{\mathrm{m},1}(z)$ and $G_{\mathrm{m},2}(z)$. The parameter estimation is done with the Levenberg–Marquardt algorithm.The regression and transfer function (RTF) model is defined as

$$y_{\mathrm{m,RTF}}(k) = G_{\mathrm{m,RTF}}(z)y_{\mathrm{m,SR}}(k - n_{\mathrm{k}})$$

with the results of the static regression model $y_{\mathrm{m,SR}}$ delayed by n_{k} samples as input for the discrete transfer function. The gain of $G_{\mathrm{m,RTF}}(z)$ is constrained to be 1 in order to add dynamic properties only. The structure is illustrated in Fig. 5.

4. Measurements and model parametrization

The acquired measurements available for model training and validation are obtained by the combustion of wooden pellets from the furnace shown in Fig. 1. They are presented in this section together with the resulting parameterization of the global and local models as well as the partitioning space and the membership functions.

4.1. Available measurements

The oxygen concentration O_2 in the flue gas is obtained by a lambda sensor and is given in %-volume/ (Nm^3) of the wet flue gas. Together with the measurement of the freeboard temperature T_{f} in K they are the most important sources of process feedback. In model based combustion control, these measurements are typically part of the underlying systems, as in [27] for a predictive controller or as used in [28] for input–output linearization. Especially in combination with a state observer they provide a valuable basis for the application of emission estimation models. A nondispersive infrared (NDIR) sensor is used to measure the concentration of carbon monoxide in the flue gas. Additional measurements were made with a combined CO-λ sensor, but have proven to be inconclusive due to high cross-sensitivities with other trace gases. The obtained CO concentration is given in ppm/(Nm^3) on dry base. The different measurements on dry and wet base are compensated by the parameters of the models [15].

To identify suitable black-box models, the data used for parameter estimation should cover the intended range of operation and reflect the process behavior. Therefore several independent open-loop measurements were taken at the test furnace in the range of 30–100% of nominal load (100 kW). The joined time series of the acquired training and validation data is presented in Fig. 6. The amplitudes are cut off at 1000 ppm for better visibility, but the cut data exhibits sensor saturation at 3400 ppm. In the bottom part of the figure the corresponding input values of O_2 and T_{f} are displayed. The data shows, that the input steps are followed by instant responses of the CO concentration and that the new setpoints are rather static than dynamic responses within the given sampling time. This is especially true for steps of the oxygen concentration, but does not change the fact that the combustion itself is a highly nonlinear and complex process. The overall amount of measured data is however limited, since not all operating conditions are represented equally well.

For the representation of carbon monoxide, the excess air ratio λ is often chosen as independent variable [22], but the oxygen concentration can be used instead as shown in Fig. 7 without introducing fundamental distortions. Based on the CO-O_2 characteristic, a furnace specific optimal level for the oxygen concentration $O_{2,\mathrm{opt}}$ can be found for which the average CO concentration has a global minimum. If the oxygen level falls below this threshold the CO emissions begin to rise steeply. Alternatively, the CO values can be displayed over the freeboard temperature as shown in Fig. 8. This allows to find a furnace specific temperature limit T_{crit}, that once underrun is followed by a

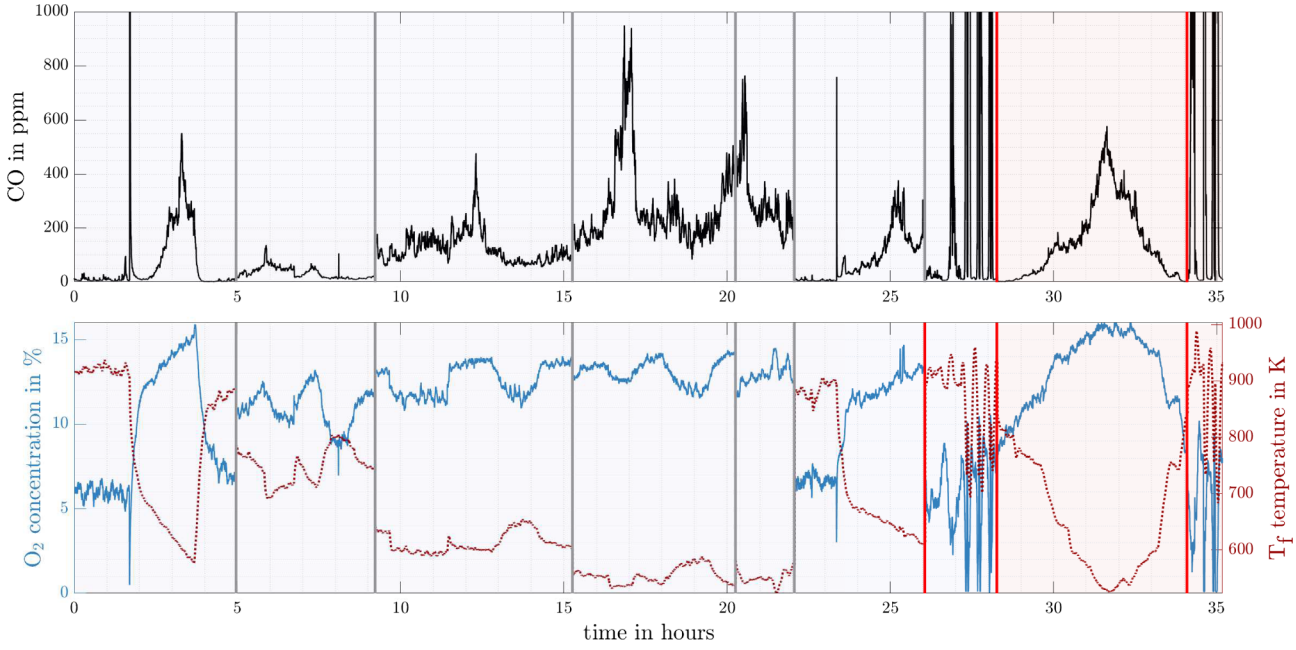


Fig. 6. Top: Available CO data for model training and validation acquired from the combustion of wooden pellets by the test furnace. The independent measurements cover the operating range and include dynamic step responses. The data used for training is separated by grey lines and the validation data by red lines. Bottom: Corresponding oxygen concentration O_2 of the flue gas and temperature of the freeboard T_{f} over the accumulated time.

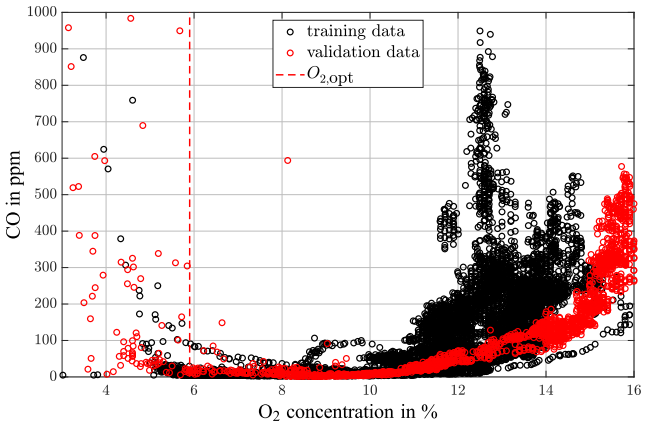


Fig. 7. Concentration of carbon monoxide over the measured oxygen concentration. The dashed line marks the furnace specific optimal O_2 level for minimal CO emissions in steady state at nominal load. The figure shows the separation into training and validation data in the CO-O_2 plane.

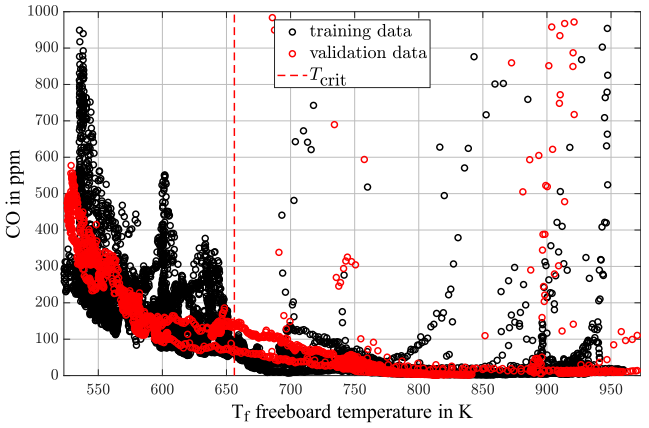


Fig. 8. Concentration of carbon monoxide over the measured freeboard temperature. The dashed line marks the beginning of significantly rising CO concentrations due to low combustion temperatures. The figure shows the separation into training and validation data in the CO-T_f plane.

considerable rise in the carbon monoxide concentration. For higher temperatures, significant changes are mainly influenced by a lack of oxygen. The exact location of these limits however cannot be determined exactly for the given measurements, but is obtained by nonlinear optimization as presented in the next section.

4.2. Partitioning

The CO-O_2-T_f data suggests to map the partitioning space with three zones, which is presented in Fig. 9. The first zone is called the oxygen starvation zone, where the temperature is high but oxygen levels are not sufficient for a complete burnout of the gaseous elements in the hot flue gas. The applicability of this zone is indicated by the oxygen concentration for minimal CO emissions from Fig. 7. The second zone is called operating zone, where the temperature is high enough for a proper burnout and sufficient oxygen supply allows for low CO emissions. The last zone is called the low temperature zone, where a complete burnout of gaseous elements is limited by low temperature, although sufficient oxygen is available, as in Fig. 8. Additionally, the data in the O_2-T_f plane of Fig. 9 are mainly aligned along a curve that reflects the steady state operation of the furnace. This is approximated with a spline of order two that either constrains O_2 to T_f or vice versa. A potential modeling approach based on this center spline is introduced in Section 4.4. The structures of the local models are defined on the basis of the split training data as presented in Fig. 9, but their final

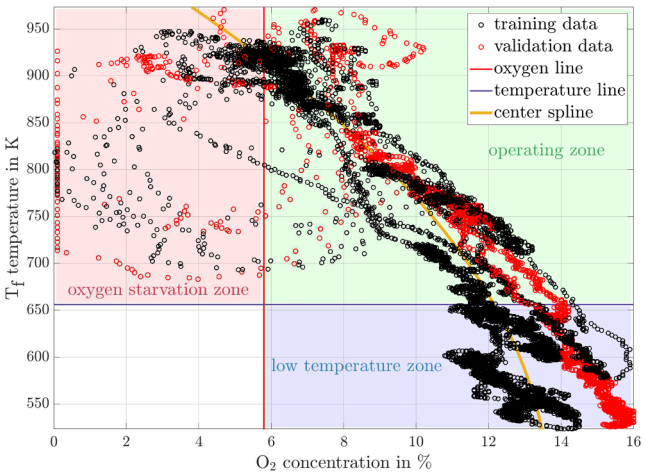


Fig. 9. Initial partitioning space for the fuzzy models with the oxygen starvation (red), operating (green) and low temperature zone (blue). The oxygen line corresponds to $O_{2,opt}$ and the temperature line to T_{crit}. Due to reoccurring CO-O_2-T_f combinations, a spline can be defined that reflects the steady state operation of the furnace.

parameters are defined after the parameterization of the membership functions.

4.3. Parametrization of the membership functions

To find suitable parameters for the Gaussian membership functions in Eq. (15) and to overcome dependencies on initial values, a genetic algorithm (GA) is applied for nonlinear optimization. The available degrees of freedom (DOF) are the centerpoints $cp_{m,p}$, the spreads $\sigma_{m,p}$ and the correlation coefficients r_m of the Gaussians. The critical temperature T_{crit} and the optimal flue gas oxygen concentration $O_{2,opt}$ as shown in Fig. 9 are made available for the optimization as well, since their exact location cannot be determined on the basis of the available measurements. The centerpoints $cp_{m,p}$ however are excluded from the optimization and determined in every iteration by fuzzy c-means clustering (FCM) [29] instead. This guarantees, that each resulting centerpoint is physically achievable by furnace operation. The resulting centerpoints can then be used for a consecutive controller design. Since the optimization is able to move the locations of T_{crit} and $O_{2,opt}$, the data assigned to the different zones varies in every iteration. Although some data points will move into other zones, the characteristics of each zone are maintained if the locations of T_{crit} and $O_{2,opt}$ are not moved significantly during the optimization. Therefore, the local model structures can be defined based on the initial partitioning space and do not have to be updated permanently. Alternatively, the model structures can be modified and adapted within the loop as well as presented in [16] for example. Because the optimization requires only a single scalar performance measure and a lot of possible model combinations exist, local SR models are selected as basis for the fuzzy model. Their application in the optimization is favorable due to the almost static input response behavior of the CO concentration as described in Section 4.1 and because static models are not affected by splitting time series data, which can lead to problems for dynamic models and clustering [30]. This results in a fast and relatively simple optimization procedure, which updates the zone limits, retrains the SR models, merges them to a fuzzy model which is evaluated on the entire training set, updates the membership parameters and repeats these steps until the optimization is finished. The applied self partitioning fuzzy work flow based on the local SR models with predefined structures is presented in Algorithm 1.

Algorithm 1. optimize $\sigma_{m,p}$, r_m, T_{crit} and $O_{2,opt}$

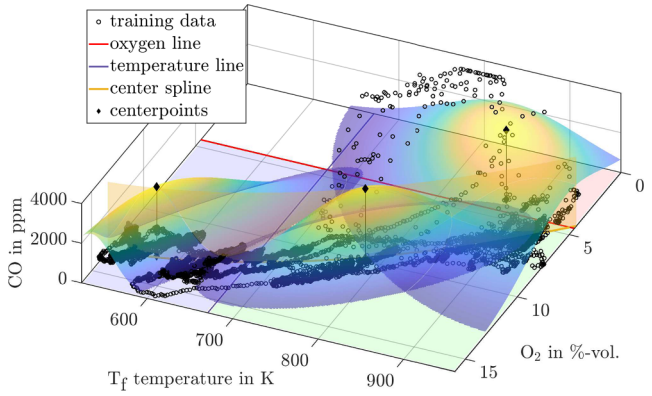


Fig. 10. CO-O_2-T_f partitioning space of the measured training data, the membership functions, the centerpoints, the center spline and the zones. The temperature and oxygen lines, the spreads and the rotation of the Gaussians are the result of the nonlinear parameter optimization, whereas the centerpoints have been defined by clustering. This figure visualizes how the fuzzy model is constructed and how the training data is assigned to the zones.

Require $y(k)$, $u(k)$, predefined local SR structures
1: initialize $\sigma_{m,p}$, r_m, T_{crit} and $O_{2,opt}$ and start GA
2: **repeat**
3: split data according to T_{crit} and $O_{2,opt}$
4: define centerpoints $cp_{m,p}$ of each zone with FCM
5: retrain the local SR models with Eq. (18) for every zone
6: merge the three local SR models with Eq. (13)
7: evaluate fuzzy SR model on entire training data → error
8: update $\sigma_{m,p}$, r_m, T_{crit} and $O_{2,opt}$
9: **untill** iter < max iter || error <ϵ
10: final $\sigma_{m,p}$, r_m, T_{crit} and $O_{2,opt}$

The obtained membership functions are therefore the same for all model combinations. It has to be noted, that this procedure may favor the results of the regression model, but that it can be conducted with other combinations of local models or additional model selection as well. Compared to an expected improvement of the model performance through additional measurements however, this can be neglected. The resulting Gaussians are illustrated in the CO-O_2-T_f space in Fig. 10, together with the new limits for T_{crit} and $O_{2,opt}$ and the center spline, which is briefly addressed in the next section. The results show, that the initial temperature and oxygen lines are shifted slightly by the optimization and that the Gaussians are rotated in the partitioning space. Additionally, centerpoints of two of the three membership functions are located almost exactly on the center spline.

4.4. Center spline model

The spline from Fig. 10 can be utilized as a new reduced system with the local coordinate l, the position along the spline. This is done by projecting the data points onto the curvature by their shortest Euclidean distance. It appears straight forward to use this CO-l plane as a new modeling space for simpler models. While the black-box structures introduced in Section 3 are depending on the inputs u_1 and u_2, models located on the spline only depend on l. The transformed and projected data is presented in Fig. 11. The data samples connected by the dashed lines are the projections of dynamic setpoint changes around nominal load and steps conducted in the oxygen starvation zone. These measurements indicate, that two very different CO levels have to be described by one value of l. Given only one independent input, a proper distinction of the different measurements is not entirely possible and is lost to some extent in the projection. Models located on the spline however achieve good performances for a stationary representation of the data with simple structures and with one input, but this approach is not pursued further.

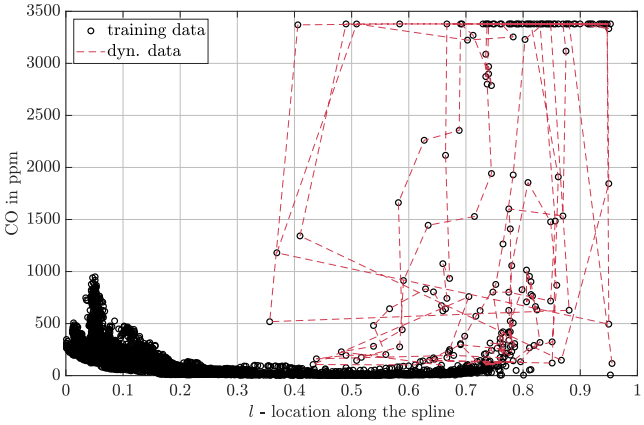


Fig. 11. Concentration of carbon monoxide over the projected location on the spline. The connected lines highlights data samples due to dynamic setpoint changes around nominal load which are difficult to describe by one input.

4.5. Black-box model parametrization

The presented data for parameter identification is sampled with Ts = 10 s. The inputs for the ARX and HM are scaled to the interval [−1 1] and the means are removed. The latter also applies to the RTF. The output is scaled to [0 3.5] for all models. The unscaled interval for the inputs O_2 and T_f is [0 16]% vol. and [530 960] K respectively. The resulting model orders are selected based on the normalized root mean squared error (NRMSE) and an evaluation criterion. Common methods are the coefficient of determination (R^2), the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) [31], of which the corrected AIC (AICc) was considered due to the small number of samples in the oxygen zone. The number of neurons for the neural networks are chosen based on the evaluations done in [15], which is a similar application.

4.5.1. Oxygen starvation zone models

The CO concentration in this zone is not affected by the temperature in such a way that it is high enough for the combustion process, but rather depends on the lack of oxygen. Although this zone is favorably not a part of steady state operation, it is close to the settings for nominal load of the furnace. Therefore transient effects and dynamic setpoint changes can have a huge impact on CO levels. The static regression model for m = 1 is obtained as

$$y_{1,SR}(k) = 9.61 - 0.72u_1(k-1) - 7.1u_2(k-1)\cdot 10^{-3}.$$

The carbon monoxide emissions in this zone heavily depend on oxygen, but the fuzzy model has a smoother transition between the zones if the temperature is considered as well. The RTF model is obtained as

$$y_{1,RTF}(k) = y_{1,SR}(k)$$

which is the SR models without alterations. The ARX model structure for the oxygen starvation zone is

$$\begin{aligned} y_{1,ARX}(k) &= 0.65y_{1,ARX}(k-1) - 0.18y_{1,ARX}(k-2) - \\ &\quad - 7.16u_1(k-1) - 2.7u_2(k-1). \end{aligned}$$

All the local Hammerstein models have piecewise linear input nonlinearities with 3 breaking points, which are depicted in Fig. 12. The accompanying transfer functions are

$$G_{1,1}(z) = \frac{1}{z^2 - 0.53z + 0.17}$$

$$G_{1,2}(z) = \frac{1}{z^2 + 0.02z + 0.41}$$

which yields the model output $y_{1,HM}(k)$ according to Fig. 4. The neural

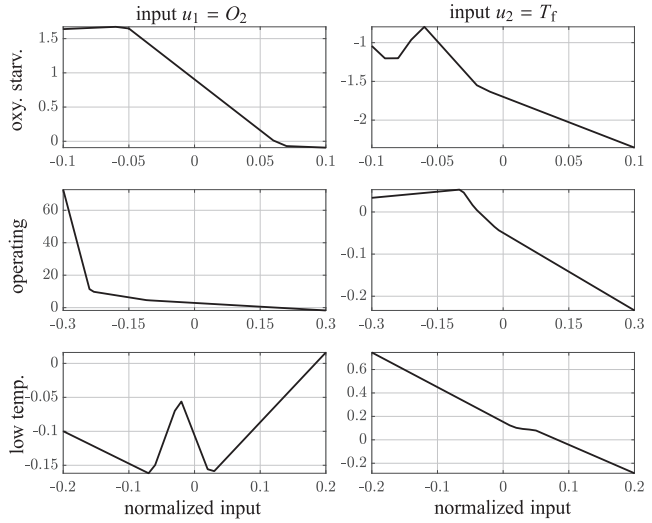


Fig. 12. Input nonlinearities of the local Hammerstein models of the three zones over the normalized inputs $u_1 = O_2$ and $u_2 = T_f$.

network has 2 neurons in the oxygen starvation zone.

4.5.2. Operating zone models

The static regression model for m = 2 is

$$y_{2,\mathrm{SR}}(k) = -0.37 - 1.80u_1(k-1) + 1.23u_1^2(k-1) + 3.25u_2(k-1) - 2.44u_2^2(k-1)$$

and the regression transfer function model to

$$y_{2,\mathrm{RTF}}(k) = -3.14y_{2,\mathrm{SR}}(k-1) + 2.90y_{2,\mathrm{SR}}(k) + 0.85y_{2,\mathrm{RTF}}(k-1).$$

The ARX model is

$$y_{2,\mathrm{ARX}}(k) = 0.56y_{2,\mathrm{ARX}}(k-1) + 0.25y_{2,\mathrm{ARX}}(k-2) - 0.019u_1(k-1) - 0.0587u_2(k-1).$$

The nonlinearities of the global Hammerstein model again with 3 breaking points are presented in Fig. 12 in the row in the middle. The transfer functions are

$$G_{2,1}(z) = \frac{z^2 - 1.05z + 0.05}{z^2 - 1.18z + 0.19}$$

$$G_{2,2}(z) = \frac{z-1}{z-0.976}.$$

The global neural network for this zone has 5 neurons in the hidden layer.

4.5.3. Low temperature zone models

The last zone has higher fluctuations of carbon monoxide than the other zones and the quality of the CO estimation in this zone is limited by the stochastic combustion conditions. The static regression model m = 3 is

$$y_{3,\mathrm{SR}}(k) = 8.65 + 0.15u_1(k-1) - 26.15u_2(k-1) + 19.73u_2^2(k-1)$$

and the accompanying RTF is

$$y_{3,\mathrm{RTF}}(k) = -1.76y_{3,\mathrm{SR}}(k-1) + 1.22y_{3,\mathrm{SR}}(k) + 0.65y_{3,\mathrm{RTF}}(k-1).$$

The ARX model is

$$y_{3,\mathrm{ARX}}(k) = 1.33y_{3,\mathrm{ARX}}(k-1) - 0.35y_{3,\mathrm{ARX}}(k-2) - 0.92u_1(k-1)\cdot 10^{-3} + 0.02u_2(k-2)\cdot 10^{-3} - 0.001u_2(k-3).$$

The nonlinearities of the Hammerstein model are presented in Fig. 12 in the bottom row and the transfer functions for this zone are

$$G_{3,1}(z) = \frac{z-1}{z-0.99}$$

$$G_{3,2}(z) = \frac{z-0.93}{z-0.92}.$$

The neural network has 3 neurons in the hidden layer of the low temperature zone.

4.5.4. Global models

The structures of the global models are in general similar to the local models, but usually of higher order due to the wider range of data that has to be modeled. The static regression model is obtained as

$$y_{\mathrm{g,SR}}(k) = 4.79 - 1.39u_1(k-1) + 0.18u_1^2(k-1) - 0.01u_1^3(k-1) + 2.54u_1^4(k-1)\cdot 10^{-4} + 7.05u_2(k-1)\cdot 10^{-4} - 1.79u_2^2(k-1)\cdot 10^{-6}$$

and the RTF model is

$$y_{\mathrm{g,RTF}}(k) = -1.11y_{\mathrm{g,SR}}(k-1) - 1.73y_{\mathrm{g,SR}}(k) + 3.85y_{\mathrm{g,RTF}}(k-1).$$

The global ARX structure is

$$y_{\mathrm{g,ARX}}(k) = 1.78y_{\mathrm{g,ARX}}(k-1) - 0.79y_{\mathrm{g,ARX}}(k-2) - 1.39u_1(k-1) + 1.38u_1(k-2) - 1.26u_2(k-1) + 1.23u_2(k-2).$$

The input nonlinearities of the global Hammerstein model are presented in Fig. 13 and the transfer functions are obtained as

$$G_{\mathrm{g},1}(z) = \frac{1}{z^2 - 0.66z + 0.08}$$

$$G_{\mathrm{g},2}(z) = \frac{1}{z^2 - 0.59z + 0.19}.$$

The neural network for this zone has 6 neurons in the hidden layer.

5. Results and validation

The obtained models are validated with the two highlighted data sets of Fig. 6 which were not available for training. The validation data covers the entire partitioning space and therefore the range of operation to verify the performance of the models in each zone, although the overall amount of data is relatively small for some zones. To provide an absolute and a relative measure, the RMSE

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

and the normalized root mean squared error (NRMSE)

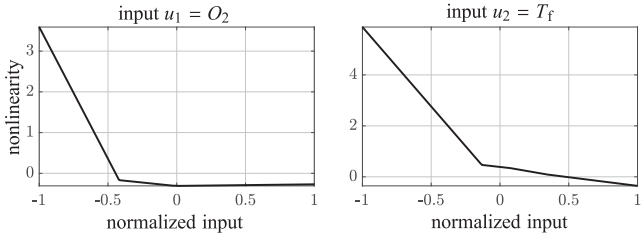


Fig. 13. Input nonlinearities of the global Hammerstein model over the normalized inputs $u_1 = O_2$ and $u_2 = T_f$.

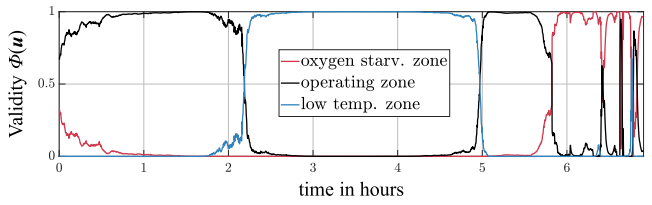


Fig. 14. Validity functions of the fuzzy model indicating how much each zone is activated during validation and how a local model contributes to the current CO concentration.

$$\mathrm{NRMSE} = 1 - \sqrt{\frac{\sum_{i=1}^{N} (y_{\mathrm{ref},i} - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_{\mathrm{ref},i} - \bar{y}_{\mathrm{ref},i})^2}}$$

are introduced, where N is the number of samples, y_{ref} is the measured reference data, $\hat{y}$ is the estimated output and $\bar{y}_{\mathrm{ref}}$ is the reference mean. Since the amplitudes in the different zones vary greatly in height, the NRMSE is used instead of the RMSE for comparison.

5.1. Local model results

The obtained local models are parametrized to perform particularly well in the assigned zone, which has to be considered in their evaluation. This is done by weighting the measured CO data available for validation as presented in Fig. 6 with the validity functions Φ_{m} instead of splitting the validation data. The transition between the zones is then already accounted for. The obtained validity functions for the inputs $\boldsymbol{u}$ are presented in Fig. 14. The contribution of the submodels to the CO concentration can now be evaluated properly. The results in Table 2 present the NRMSE of the different local model structures. In the oxygen starvation zone (1) and the temperature zone (3) all structures perform equally well. The parameters converge fast due to the distinct process dynamics, which are particularly expressed by the oxygen concentration. In the operating zone (2) the lowest performances are located. This can be addressed for example by penalizing the model performances in the different zones in the optimization. Typically the SR and the ARX model are the best performers. Table 2 can now be used to find the composition for the fuzzy model by selecting the best model of each zone. This direct selection is possible, because the validation data is weighted with the validity functions Φ and therefore already

Table 2
Overview of the local model performance on the weighted validation data evaluated with the NRMSE in %.

Local model validation					
Zone/ Φ	SR	RTF	NN	ARX	HM
1	74.8	74.8	67.9	81.1	74.1
2	45.8	45.6	21.9	26.4	27.5
3	73.5	73.6	53.7	60.3	57.3

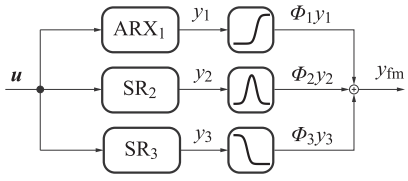


Fig. 15. Structure of the best performing fuzzy model. The selection of the local submodels is done according to the results presented in Table 2. The best performance is achieved by an ARX in the first, a SR in the second and a SR model in. the third zone.

Table 3
Results of the global model performances on the validation data compared to fuzzy models with substructures of the same type, all evaluated with the NRMSE in %.

Overall model validation						
Type	SR	RTF	NN	ARX	HM	BF
Global	64.8	68.0	68.4	22.5	69.2	–
Fuzzy	70.1	69.1	58.0	65.8	63.5	74.1

considers the effects of the fuzzification. The best fuzzy (BF) model is obtained by combining an ARX for the first, a SR for the second and a SR structure for the third zone, which is presented in Fig. 15 in block form.

5.2. Global model results

For a fair comparison of the fuzzy and the global models it is necessary to consider how the validation data is distributed in the local zones. If this is not considered, the validation can be distorted by the selected data. This can be addressed by weighting or normalizing the performance with the duration a zone is represented or by other measures, like rating the relevance of a zone for the plant operation for example. The resulting NRMSE's of the fuzzy and the global black-box models have therefore been weighted by the duration a zone is active and are presented in Table 3. For comparability, all fuzzy models only contain local structures of the same type (e.g. all three are ARX), except for the best fuzzy model. The results show, that the additional effort to separate the operating range into distinct zones and to mix different substructures is of advantage. The multivariate static regression model is apparently a suitable structure, as was also concluded in [15]. However, the extension to the RTF is not associated with major improvements. The Hammerstein model is an effective approach as well, especially for a global application with simple input nonlinearities. Of the investigated black-box structures, the global ARX model is found to be the most unfit for the estimation of carbon monoxide for the given furnace data. Local ARX models can achieve good results however, which is also shown in [32] for a radial basis function ARX structure for the prediction of NO_x emissions in thermal power plants.

The results of the best combined fuzzy model are visualized in Fig. 16 together with the fuzzy ARX, the global neural network and the global Hammerstein model in the upper plot. Due to the different amplitudes, the validation series is split. The middle plot shows the residuals together with the RMSE, where unadressed process dynamics of some models become visible. Although further inputs or more data may solve this issue, a closer look at the accompanying temperature in Fig. 6 indicates that the combustion is close to being extinguished, which is an operating condition that should be avoided. The model outputs in Fig. 16 have been cut off according to the sensor saturations, although the models are able to extrapolate the CO concentration.

The results of the analytical approach are shown in Fig. 17. The kinetic model is capable to estimate the CO concentration with all the introduced simplifications. However, compared to most black-box models and the effort taken to obtain the modeling equations, the performance is questionable. Some of the difficulties of this approach are the Arrhenius coefficients which are hard to obtain. Given these results, black-box modeling is the more efficient approach, which is also reflected in literature by the various applications of data driven models for the estimation of NO_x, as addressed in the introduction.

6. Discussion

Of all the presented methods to find suitable models for the estimation of CO emissions of a small-scale biomass combustion furnace a fuzzy model shows the most promising results. It can provide the

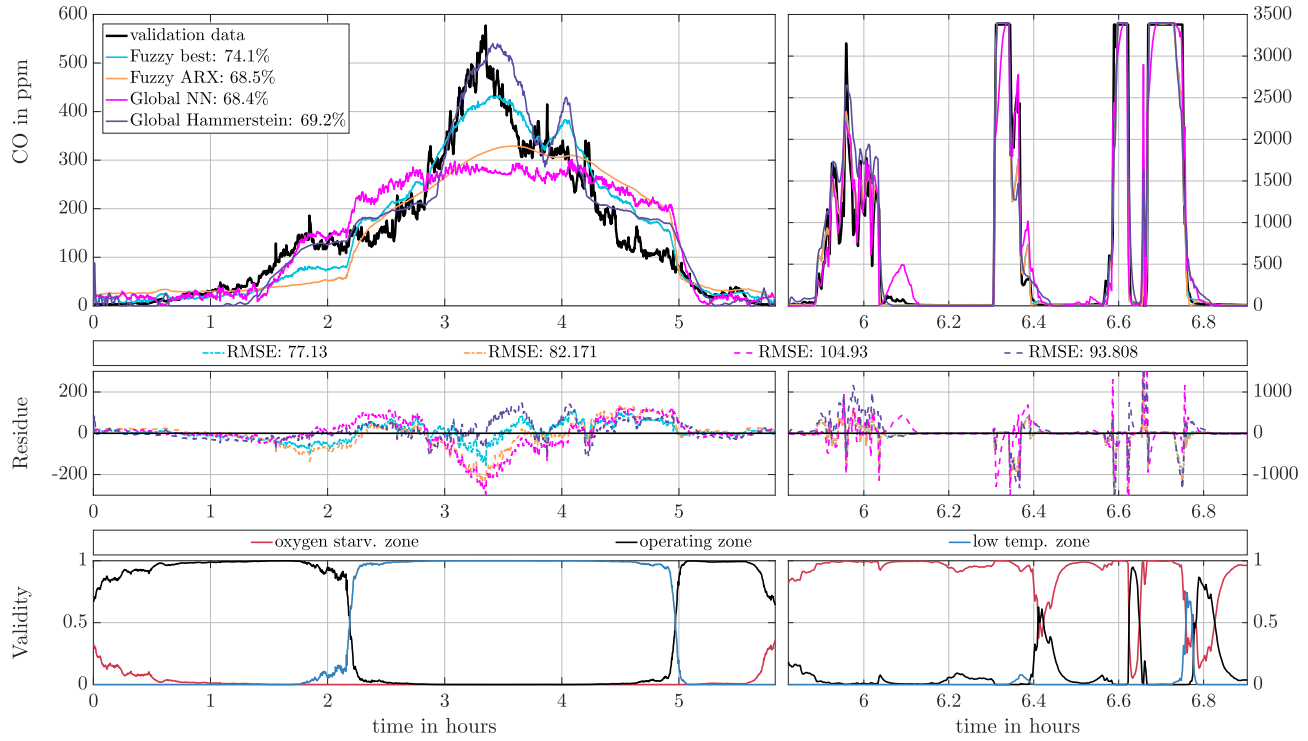


Fig. 16. Top: Comparison of the best combined fuzzy model, the fuzzy ARX, the global NN and the global HM compared to the validation data. The fit is depicted by the weighted NRMSE presented in Table 3. Middle: Residuals of the models in ppm with the weighted RMSE. Bottom: Validity functions of the fuzzy model.

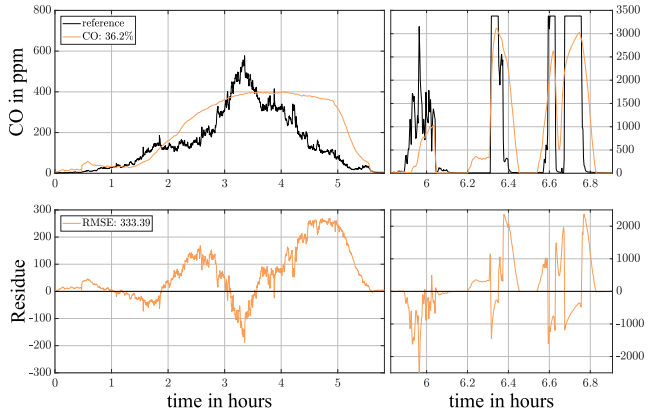


Fig. 17. Top: NRMSE of the analytical model. Bottom: Residual in ppm and RMSE.

required simplicity which is necessary for an implementation in a model based controller or as a soft sensor for the given limited computational hardware of the plant. It is also able to cover the entire operating range, although the performance for low freeboard temperatures could be improved. The training framework can be extended for example with a cross-validation for structure selection in the optimization loop. The obtained models can be implemented into a predictive controller to consider the emission limitations as constraints for example or for fault detection and isolation of unreliable CO sensors.

The following contains a listing of the advantages and disadvantages of the different model types introduced in this work. Whenever the modeling effort is addressed, it is implicitly assumed that the same amount of data preprocessing has to be done for all methods.

Fuzzy and data driven models:

A major advantage of the fuzzy model is that any structure can be used for the different subsections which allows to select the best mix of structures. This is however accompanied by the additional a priori effort to define and parametrize the membership functions, as well as potential problems caused by unfavorable splits of the zones. However, the large differences in the CO amplitudes should be taken into account and treated separately.

- SR: A big advantage of the static regression approach is the easy handling and the training with a least squares algorithm. The best overall results for the estimation of CO have been achieved with this structure. For an implementation, pre-filtering the oxygen concentration with a low order filter can be considered to improve steady state estimation.
- RTF: The results of the regression transfer function model show, that no significant improvements compared to the SR have been achieved. Although its a simple add-on to an existing regression model, the added value is not really worth the effort.
- ARX: The global ARX model is found to be the most unfit structure for the estimation of carbon monoxide for the investigated furnace. However, the local ARX structures in combination with the fuzzy framework are able to estimate the CO concentrations very well.
- HM: The application of Hammerstein models to estimate carbon monoxide emissions is not common, at least the authors could not find any references. The general applicability of the structure is however very conceivable as demonstrated in this work. By choosing a piecewise linear function for the static nonlinearity, the overall effort can be kept small and the structure is simple.
- NN: The neural network is straightforward to estimate with the right tools and can quickly result in models with a good performance. This makes it the first choice, if fast and efficient results are required.

Kinetic model:

The kinetic reaction model is difficult to obtain, has a large number of unknown parameters and due to the Arrhenius expressions, is numerically stiff. Although it is possible to find sets of parameters after several nonlinear optimization runs, other models yield better results and are more efficient.

7. Conclusion

Several different approaches are presented in this work to find simple and capable structures for the estimation of carbon monoxide for the investigated small-scale biomass combustion plant for wooden pellets. The focus of this work was to obtain models which can be used by furnaces with limited computational power, either for controller design and/or fault detection of carbon monoxide sensors. Of the investigated approaches, a fuzzy model based on different local structures is considered to be the most suitable regarding complexity and performance. The best performing local structures are a multivariate regression model and an autoregressive model with exogenous inputs. The Hammerstein model presents itself as a capable approach as well, especially for a global description. Without question, neural networks have proven their capability to estimate the emissions released from different kind of combustion processes and other applications. The kinetic process model however achieved inferior results compared to the black-box models and is additionally more complex to obtain.

The presented fuzzy framework is based on the distinction between three zones for the classification of the data in the introduced partitioning space. Separating the data as presented is based on literature and observations and can be reproduced for different furnaces. It may seem that the effort to obtain the fuzzy models is high compared to what little effort is needed for a global neural network given the right tools. This is however compensated by the performance advantage and the fact that a fuzzy regression model for example is simple enough for the intended operation in small-scale furnaces, shows good overall results and is efficiently parametrized in the introduced training algorithm, which yields a fully parametrized fuzzy model.

The next steps are to investigate the performance of the fuzzy model in combination with a controller or as a soft sensor, potentially extended with adaptive parameter estimation and a state observer. Further improvements for all structures may be achieved by restricting the carbon monoxide levels to the governmental limitations of the applied country or to use methods like deep learning for parameter estimation.

Acknowledgments

The authors want to thank the partners at the FH-Burgen-land, Herz Energietechnik GmbH and Binder Energietechnik GmbH for their continuing support and the supply with test-data of the furnace. This project (#853.606) is funded by the Climate and Energy Fund and conducted under the 2015 Energy Research Program.

References

[1] Ikonen E, Najim K, Kortela U. Neuro-fuzzy modelling of power plant flue-gas emissions. Eng Appl Artif Intell 2000;13:705–17.
[2] Koppejan J, Van Loo S. The handbook of biomass combustion and co-firing. Routledge; 2012.
[3] Korpela T, Björkqvist T, Majanne Y, Lautala P. Online monitoring of flue gas emissions in power plants having multiple fuels. IFAC Proc Vol 2014;47:1355–60.
[4] Gambarotta A, Morini M, Zubani A. A non-stoichiometric equilibrium model for the simulation of the biomass gasification process. Appl Energy 2018;227:119–27.
[5] Zhao H, Shen J, Li Y, Bentsman J. Coal-fired utility boiler modelling for advanced economical low-nox combustion controller design. Control Eng Practice 2017;58:127–41.
[6] Haseli Y, Van Oijen J, De Goey L. Modeling biomass particle pyrolysis with temperature-dependent heat of reactions. J Anal Appl Pyrol 2011;90:140–54.
[7] Privara S, Cigler J, Váňa Z, Oldewurtel F, Sagerschnig C, Žáčeková E. Building modeling as a crucial part for building predictive control. Energy Build 2013;56:8–22.
[8] Nussbaumer T. Combustion and co-combustion of biomass: fundamentals, technologies, and primary measures for emission reduction. Energy Fuels 2003;17:1510–21.
[9] Carvalho L, Wopienka E, Pointner C, Lundgren J, Verma VK, Haslinger W, et al. Performance of a pellet boiler fired with agricultural fuels. Appl Energy 2013;104:286–96.
[10] Varol M, Atimtay AT, Olgun H, Atakül H. Emission characteristics of co-combustion of a low calorie and high sulfur–lignite coal and woodchips in a circulating fluidized bed combustor: Part 1. Effect of excess air ratio. Fuel 2014;117:792–800.
[11] Fujii S, Tomiyama S, Nogami Y, Shirai M, Ase H, Yokoyama T. Fuzzy combustion control for reducing both co and nox from flue gas of refuse incineration furnace. JSME Int J Ser C Mech Syst, Machine Elements Manuf 1997;40:279–84.
[12] Korpela TM, Björkqvist TK, Lautala PA. Control strategy for small-scale wood chip combustion. IFAC Proc Vol 2009;42:119–24.
[13] Song J, Romero CE, Yao Z, He B. A globally enhanced general regression neural network for on-line multiple emissions prediction of utility boiler. Knowl-Based Syst 2017;118:4–14.
[14] Li N, Lu G, Li X, Yan Y. Prediction of nox emissions from a biomass fired combustion process based on flame radical imaging and deep learning techniques. Combust Sci Technol 2016;188:233–46.
[15] Korpela T, Kumpulainen P, Majanne Y, Häyrinen A, Lautala P. Indirect nox emission monitoring in natural gas fired boilers. Control Eng Practice 2017;65:11–25.
[16] Smrekar J, Potočnik P, Senegačnik A. Multi-step-ahead prediction of nox emissions for a coal-based boiler. Appl Energy 2013;106:89–99.
[17] Chong A, Wilcox S, Ward J. Prediction of gaseous emissions from a chain grate stoker boiler using neural networks of arx structure. IEE Proc-Sci, Measur Technol 2001;148:95–102.
[18] Kalogirou SA. Artificial intelligence for the modeling and control of combustion processes: a review. Prog Energy Combust Sci 2003;29:515–66.
[19] Caposciutti G, Barontini F, Antonelli M, Tognotti L, Desideri U. Experimental investigation on the air excess and air displacement influence on early stage and complete combustion gaseous emissions of a small scale fixed bed biomass boiler. Appl Energy 2018;216:576–87.
[20] Kortela J, Jämsä-Jounela S-L. Model predictive control utilizing fuel and moisture soft-sensors for the biopower 5 combined heat and power (chp) plant. Appl Energy 2014;131:189–200.
[21] Di Blasi C. Modeling chemical and physical processes of wood and biomass pyrolysis. Progress Energy Combust Sci 2008;34:47–90.
[22] Kaltschmidt M, Hartmann H, Hofbauer H. Energie aus biomasse, Grundlagen, Techniken und Verfahren. Berlin-Heidelberg: Springer-Verlag; 2009.
[23] Foresee FD, Hagan MT. Gauss-newton approximation to bayesian learning, in. Proceedings of the 1997 international joint conference on neural networks, vol. 3. Piscataway: IEEE; 1997. p. 1930–5.
[24] Nelles O. Nonlinear system identification: from classical approaches to neural networks and fuzzy models. Springer Science & Business Media; 2013.
[25] Billings SA. Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains. John Wiley & Sons; 2013.
[26] Ljung L. System identification: Theory for the User. Prentice Hall PTR; 2006.
[27] Kortela J, Jämsä-Jounela S. Modeling and model predictive control of the biopower combined heat and power (chp) plant. Int J Electr Power Energy Syst 2015;65:453–62.
[28] Gölles M, Reiter S, Brunner T, Dourdoumas N, Obernberger I. Model based control of a small-scale biomass boiler. Control Eng Practice 2014;22:94–102.
[29] Bezdek JC. Pattern recognition with fuzzy objective function algorithms. Springer Science & Business Media; 2013.
[30] Liao TW. Clustering of time series data – a survey. Pattern Recog 2005;38:1857–74.
[31] Samad MFA, Nasir ARM. Comparison of information criterion on identification of discrete-time dynamic system. Int J Eng Technol 2018.
[32] Peng H, Ozaki T, Toyoda Y, Shioya H, Nakano K, Haggan-Ozaki V, et al. Rbf-arx model-based nonlinear system modeling and predictive control with application to a nox decomposition process. Control Eng Practice 2004;12:191–203.

Glossary

AIC: Akaike Information Criterion.
AICc: corrected Akaike Information Criterion.
ARX: AutoRegressive model with eXogenous inputs.
BF: Best Fuzzy.
BIC: Bayesian Information Criterion.
DOF: Degree of Freedom.
FCM: Fuzzy C-Means.
GA: Genetic Algorithm.
HM: Hammerstein Model.
MLP: Multi-Layer Perceptron.
MPC: Model Predictive Controller.
NARX: Nonlinear ARX.
NDIR: NonDispersive InfraRed (Sensor).
NN: Neural Network.
NRMSE: Normalized Root Mean Squared Error.
ppm: parts per million.
RMSE: Root Mean Squared Error.
RTF: Regression and Transfer Function.
SR: Static Regression.