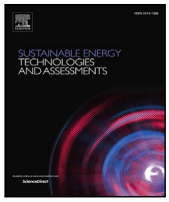
Sustainable Energy Technologies and Assessments 58 (2023) 103321

Contents lists available at ScienceDirect

Sustainable Energy Technologies and Assessments

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



Simplified finite volume-based dynamic modeling, experimental validation, and data-driven simulation of a fire-tube hot-water boiler

Marco Tognoli [a], Shayan Keyvanmajd [b], Behzad Najafi [a,*], Fabio Rinaldi [a]

[a] Dipartimento di Energia, Politecnico di Milano, Via Lambruschini 4, Milano 20156, Italy
[b] Department of the Built Environment, Eindhoven University of Technology, 5600 MB Eindhoven, Netherlands

ARTICLE INFO

Keywords:
Fire-tube boiler
Dynamic modeling
Data-driven modeling
Experimental validation
Thermal performance simulation

ABSTRACT

This paper proposes and implements a novel approach to simulate the dynamic behavior of hot-water fire-tube boilers, in which physical phenomena-based and data-driven modeling methodologies have both been employed. The first model, which includes a reduced one-dimensional finite volume method to simulate the flue-gas side's behavior, can be employed for accurate sizing of the unit considering end-users' dynamic consumption profile. In the data-driven model instead, machine learning algorithms are used to estimate hot water's supply temperature, which makes it a suitable tool for real-time prediction and model predictive control. Utilizing the latter model in combination with an advanced management system allows reducing the plant's energy consumption and enhancing its controllability. Employing the measurements performed in an Italian industrial firm, the developed model is validated and is demonstrated to have a limited thermal efficiency estimation bias of 1.2%. Furthermore, the data-driven model achieves a mean absolute relative difference error lower than 6%, demonstrating its acceptable accuracy for various prediction horizons.

Introduction

The hot water demand with low to medium required volumetric flow rates, for both residential and industrial sectors, is commonly addressed by fire-tube boilers [1]. Optimal sizing of these units, while considering the end-users' consumption profile, is a crucial step to evade the utilization of unnecessarily oversized units and reduce the corresponding required capital cost. However, such a sizing procedure (specifically for notably varying demand profiles) requires a detailed and accurate model that can simulate the boiler's dynamic behavior [2,3]. Several previous studies have investigated the thermal behavior of water-tube boilers that are commonly utilized for supplying steam [4–12], a few of which have also modeled the heat transfer rate of flue-gases [13,14]. Fire-tube boilers, instead can either be configured to supply medium-to-low saturated steam or sub-cooled water. Dynamic thermal modeling of these units, have mainly been dedicated to fire-tube boilers utilized for saturated steam generation. A few studies have employed Computational Fluid Dynamics (CFD) models [15–18] that describe the underlying physical phenomena with a promising accuracy, however, utilizing them results in a significant computational cost.

Simplified finite volume modeling approach is an alternative methodology that has been employed in this area in order to achieve an acceptable accuracy, while reducing the computational cost [19]. Sorensen et al. [20] implemented a dynamic thermal model for the traditional fire-tube boilers supplying saturated steam. The authors adopted a sub-modeling approach to simulate the behavior of the flue-gas side. A more detailed approach has been proposed by Gutierrez Ortiz [21], utilizing a one-dimensional finite volume methodology to describe the flue-gas side's heat transfer mechanism. The latter modeling methodology has also been employed in previous studies, conducted by Tognoli et al. [2,3] focused on dynamic thermal behavior modeling and optimal sizing of fire-tube boilers used for supplying vapor at low-medium pressure. In these studies, the authors first simulated the radiative thermal heat transfer phenomena of the flame in the combustion chamber employing a gray-gas model, as proposed by Hottel–Sarofim [22], with a net heat transfer contribution towards the assumed gray-surface surrounding walls; the Dittus–Boelter correlation [23] was instead utilized for the internal flow. At the water-side, the heat transfer between the hot metal walls and the saturated water was described by the Cooper nucleate pool boiling correlation [24].

* Corresponding author.
E-mail address: behzad.najafi@polimi.it (B. Najafi).

https://doi.org/10.1016/j.seta.2023.103321
Received 15 December 2022; Received in revised form 9 June 2023; Accepted 10 June 2023
Available online 23 June 2023

Nomenclature

c_P	isobaric specific heat coefficient kJ/°Ckg
c_V	isocoric specific heat coefficient kJ/°Ckg
d	diameter m
e	helical coil's wire thickness m
f	Fanning friction factor
g	equivalent radiative thermal resistance W/m^2K^4
h	specific enthalpy kJ/kg
h_{conv}	convective heat transfer coefficient W/m^2K
k	linear thermal conduction coefficient W/mK
$\overline{k}_t$	average metal conductive heat transfer coefficient kW/m°C
L	length m
LHV	fuel mass lower heating value MJ/kg
m	mass kg
MAE	Mean Absolute Error
$MARD$	Mean Absolute Relative Difference %
MSE	Mean Squared Error
N	number of points/discrete finite volumes
$n_{t,j}$	number of tubes in a given gas pass section at given tube banks
Nu	Nusselt number
P	Pressure bar
p	helical coil's pitchm
Pr	Prandtl number
$\dot{Q}$	heat transfer rate kW_t
$\dot{q}$	specific heat transfer rate by unit area kW_t/m^2
R	radius m
R^2	Coefficient of determination
Ra	Rayleigh number
Re	Reynolds number
T	temperature °C
t	time s
T_s	simulation time s
V	volume m^3
Y	fuel axial heat release fraction MJ/MJ_{tot}
x	axial coordinate m
o.r.	of reading

Greek

ϵ	residuals
η	first law efficiency term
ρ	density kg/m^3
σ	Stefan Boltzmann's constant
ε	emissivity

Subscripts

$2nd$	2nd gas pass
$3rd$	3rd gas pass
CC	combustion chamber
$conv$	convective
f	feedwater
$fuel$	fuel
g	flue gas
h	hydraulic
i	finite element index
in	inlet
j	tube bank of reference
out	outlet condition
rad	radiative
ref	reference temperature conditions (0 K)
t	metal
$turb$	turbulators
$wall$	tube wall

An experimental research activity performed by Morelli et al. [19] has investigated the convective and radiative heat transfer mechanism between the hot methane-exhaust gases and the surrounding tube's walls by applying a reduced Finite Volume Methodology (FVM) at the flue-gas side. In this study, based on the obtained results, an optimized convective heat transfer correlation is proposed, while a gray-gas model employing the weighted sum of grey gas coefficients (WSGG) (as proposed by Smith et al. [25]) describes the radiative heat flux. A similar procedure has been investigated by the work developed by Abene et al. [26] in the context of the start-up process of a 3 gas-passes fire-tube boiler. In this work, a detailed model of the water sub-cooling region is also included, while the water-side convective heat transfer mechanism is described employing the Churchill–Chu [27] correlation of external natural convection in horizontal cylinders (motivated by the quasi-stationary forced movement of water inside the boiler vessel). Although the mentioned works describe the dynamic thermal performance of steam-supplying fire-tube boilers with acceptable accuracy, no previous comprehensive work has been conducted on developing an accurate model for hot-water boiler architecture.

Helical-coil turbulators are one of the commonly employed metallic inserts that are utilized for enhancing the heat transfer in fire tubes. Turbulators promote the convective heat transfer rate by enhancing the flow turbulence pattern. The resulting benefits, in terms of enhancement in the convective heat transfer coefficient (compared with the net heat transfer rate of a traditional smooth tube), have been analyzed by Ji et al. [28] over a wide range of operating conditions and insert geometries. The authors also observed a drastic increase in friction losses motivated by the enhanced turbulence caused by the inserts. In the boiler that is considered in the present investigation, helical-coil type inserts are present, the geometry of which is described by the outer coil diameter, the pitch distance, and the metal wire diameter. Several works assessed the heat transfer performances and friction losses of helical-coil inserts [29–32] while other authors performed experimental investigations at a wide range of operating conditions employing distilled water and propylene glycol [33], plain liquid water flow [34], and externally heated air [35]. For the scenario that is considered in the present study, the heat exchange phenomena of medium to cold flow of exhaust gases, which exchange heat with the colder metal walls, is described in details in the study conducted by Tognoli et al. [36]. In this study, the authors experimentally investigated the ensemble radiative and convective heat transfer, while also analyzing the friction factor, of internally heated tubes. Accordingly, the authors proposed parametric Nusselt and Fanning friction factor correlations that model the effect of turbulator's geometry.

On the other hand, a promising approach that can be employed to reduce energy consumption and enhance the controllability of the fire-tube boilers is improving the corresponding control (e.g. employing model predictive control) and setpoint management strategy [37] of these units. Implementation of such strategies requires a robust and rapid model that permits near-real-time behavior prediction. Accordingly, the significant calculation cost of the previously described physical phenomena-based model makes it unsuitable for such applications. Furthermore, this modeling approach requires boiler units' detailed geometrical data and material characteristics, while, in many industrial cases, the detailed geometrical configuration of the boiler

is not provided by the third party (e.g. manufacturer). The physical phenomena-based model is thus not a proper choice in such case studies (in which such data is unavailable). In order to address these issues, employing data-driven modeling methodology [38] is a viable alternative, which has successfully been utilized in the literature for modeling complex phenomena in heat transfer [39] and fluid mechanics [40] Further, the application of a data-driven model for predicting the dynamic behavior of the system, is of major interest in the presence of low data granularity [41]. Among the available data-driven model architectures, the Artificial Neural Networks (ANNs) category has been utilized by Romeo et al. [42] aiming at describing the thermal performance of a biomass-fed boiler under progressive fouling phenomena. The methodology that was followed in this study was driven by a physical approach aiming at decomposing the boiler assembly into simpler sub-sections, which resulted in promising model accuracy. In this context, a relevant investigation on data-driven modeling of a power plant output has been conducted by Smrekar et al. [43], modeling the plant's output power. Separate models were developed for the boiler and the turbine sections employing ANNs, which were later trained using the plant's data that was measured on-site. However, no previous comprehensive work has investigated the overall boiler's (and specifically a fire-tube unit's) dynamical thermal performances employing a data-driven modeling approach.

Accordingly, considering the above-mentioned motivations, in the present work, dynamic thermal models for a hot-water fire-tube boiler are developed. As the first step, employing a simplified 1-D finite volume methodology, a physical model is developed. Next, the experimental data, obtained through performing measurements on the considered case study, is utilized to validate the latter developed model. In the following step, in order to control the fuel flow rate supplied to the burner, provided the desired hot-water temperature setpoint (which dynamic behavior is affected by the unsteady hot-water request), a PID (Proportional-Integral-Derivative) controller is included within the model. The boiler cumulative thermal efficiency, addressing the measured thermal demand, is determined, which is estimated beginning from the measured natural gas consumption profile. Furthermore, as secondary indicators, the deviation of the hot-water supply temperature and natural gas flow rate from the corresponding experimental measurements are estimated. It is worth mentioning that a similar modeling approach has been employed by the authors in the dynamic modeling of saturated steam supplying fire-tube boilers [2,3]. Though, in the present work, more detailed convective heat-transfer correlations [19, 36] and flue-gas radiative heat transfer [25] mechanisms have been considered. Furthermore, the presence of sub-cooled water in presence of natural convection requires the application of suitable heat transfer equations. Moreover, dynamic thermal behavior modeling of hot-water boilers has not been investigated by the authors in their previous works.

Alongside the mentioned steps, data-driven predictive thermal models are developed aiming at describing the boiler's dynamic thermal behavior (as a whole) for a range of prediction horizons. In the corresponding methodology, state-of-the-art machine learning(ML)-based pipelines are first implemented. Next, for each pipeline, a genetic algorithm-based [44,45] pipeline optimization procedure is conducted to determine the algorithm (and the corresponding tuning parameters) [46], using which the highest performance (estimation accuracy) is achieved. Finally, the obtained estimation errors are compared and correspondingly reported.

In summary, the novel contributions of the present study can be summarized as:

- Proposing and implementing a physical-based dynamic model of a hot-water fire-tube boiler, in which a simplified finite-volume modeling approach is employed at the flue-gas side's behavior. The developed model can be employed for the accurate sizing of boiler units while considering the end-user's dynamic demand profile.
- Developing a data-driven model for predicting the hot-water supply temperature. The developed data-driven model allows a near real-time estimation of the boiler unit's dynamic behavior, which permits reducing the required computational resources (compared to a more resource-expensive physical phenomena-based model), making it a more suitable choice for model predictive control applications.

Description of the considered fire-tube boiler

In the present section, a general overview of the operation principle of the considered fire-tube boiler (utilized for supplying hot-water) is presented.

Boiler description

The operation principle of a fire-tube boiler, employed for generating hot-water, is based on the heat exchange between the hot flue gases (resulting from the combustion process) and the surrounding sub-cooled water. The combustion process is controlled by the burner, which manages the air and fuel flow rates. The control scheme is operated by a PID controller, the objective of which is stabilizing the water temperature inside the boiler close to the desired setpoint. In the considered boiler model, a premixed natural-gas fed burner allows the development of the flame along the combustion chamber [47]. The chemical species that are undergoing the combustion process exchange heat (through radiation and convection) with the surrounding combustion chamber's wall [48], which in turn exchanges heat with the colder water mass that is stored in the boiler vessel. The flue gases, that are still hot, subsequently stream into the 2nd gas pass, which is a single smooth horizontal tube that enables further transfer of heat from the flue gases to the water side. Finally, the remaining energy content of flue-gases, is recovered at the 3rd gas pass that consists of a bundle of tubes in which helical coil turbulators (in order to enhance the convective heat transfer rate at medium-to-low temperature range) are employed.

On the other side, the sub-cooled water is progressively heated as the result of the heat flux from the combustion chamber and the gas passes. Fig. 1 demonstrates a detailed schematic representation of the considered unit.

Description of the mathematical model

In the present section, a thorough description of the methodology that has been utilized in the reduced 1-D FVM physical modeling and the predictive data-driven simulation (motivated by the corresponding considerations and assumptions) has been provided.

Mathematical model

In order to accurately model the underlying physical phenomena that occur in the fire-tube boiler, a sub-modeling approach has been utilized and separate models have been developed for simulating the flue-gas and sub-cooled water sides. A similar procedure has been carried out by the authors in the context of saturated steam supplying fire-tube boiler models [2,3]. It is worth mentioning that, in the present work, more detailed convective heat-transfer correlations [19,36] and flue-gas radiative heat transfer [25] mechanisms have been considered. Furthermore, the presence of sub-cooled water in presence of natural convection requires the application of suitable heat transfer equations. Accordingly, the following sections have been dedicated to providing a description about the approach that has been utilized for simulating each of these sides.

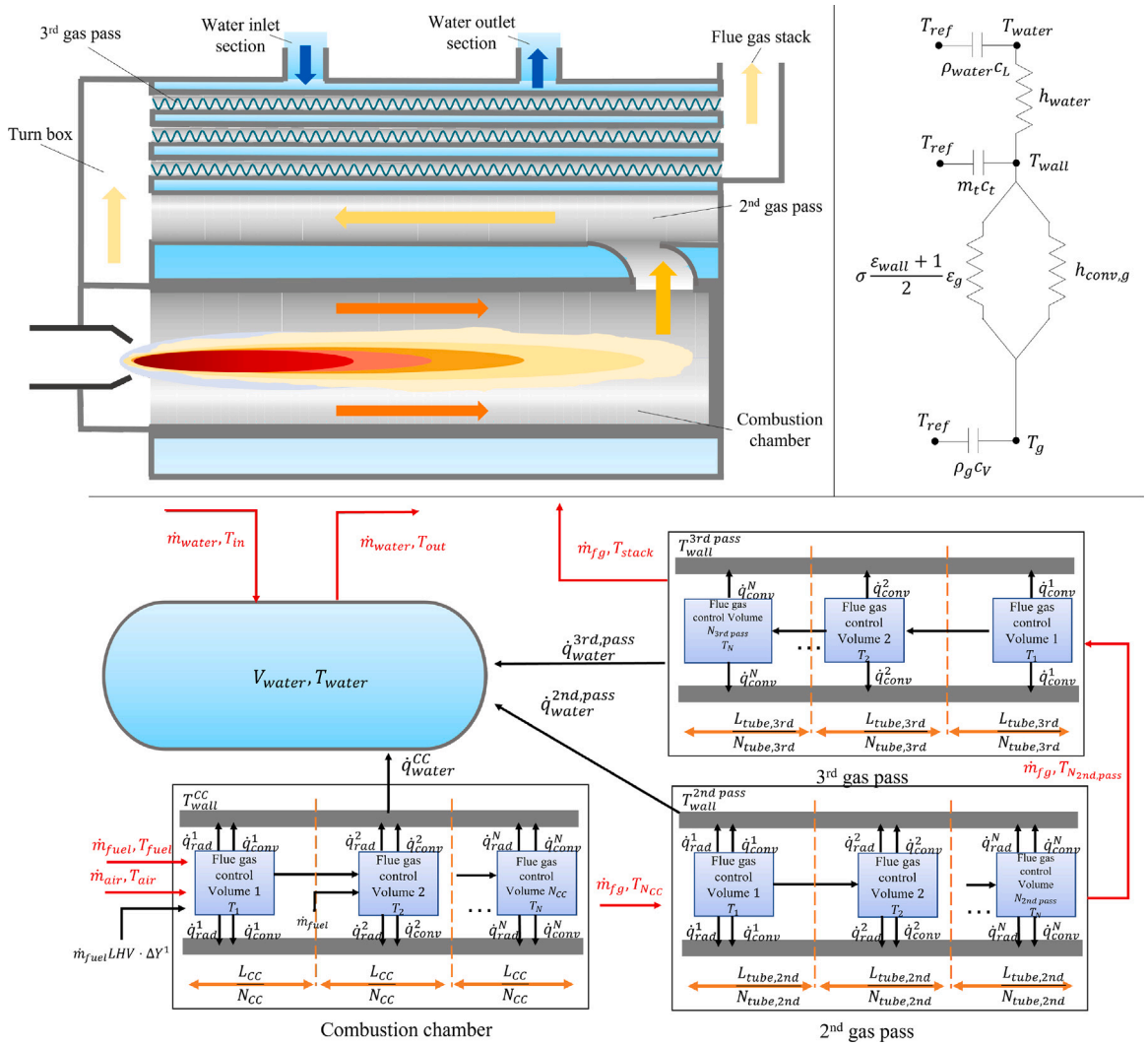


Fig. 1. (top-left) Schematic representation of the water-side and the flue gas's stream in a hot-water fire-tube boiler. (top-right) thermal–electrical analogy diagram extended from the flue-gas bulk flow condition to the water-side temperature. (bottom) The employed finite volume method modeling architecture.

Flue-gas side

The flue-gas side modeling phase is aimed at simulating the heat transfer mechanisms that take place due to the temperature gradient between the hot flue-gases and the colder tubes metal surfaces. Motivated by the different geometrical configurations and thermal behavior of various heat transfer surfaces, separate control volumes are considered in order to estimate the dynamic behavior of the flue gases in the combustion chamber and the subsequent gas passes.

Air and natural gas are pre-mixed and are later ignited, allowing the development of a longitudinal flame front along the main combustion chamber. The combustion process is assumed to be complete and instantaneous, while an exponential heat release curve is considered (as proposed by Gutierrez Ortiz [21]) The exponential behavior of the heat release function, motivated by the design of burners, describes a large portion of the heat release, due to the combustion process, located far from the injection point. The behavior of the latter combustion products is assumed to be ideal; thus, an ideal gas behavior is considered. The flame radiative heat transfer term is expressed by the flame characteristic emissivity value. It is also worth highlighting that although the air-to-fuel ratio is constant, the fuel flow rate, managed by the PID controller, is determined time-by-time by the heat balance equations discussed in the following section. The mentioned controller aims at supplying the hot-water at the desired setpoint temperature, which necessary fuel quantity is analytically determined when iteratively solving the DOEs (Discrete Ordinate Equations). Furthermore, motivated by the negligible relative effect of the frontal area compared to the more relevant lateral surfaces, for the radiative heat transfer mechanism, no axial component is considered. Flue gas properties are estimated locally employing a reduced 1-D FVM approach. The FVM discretization is performed separately at the combustion chamber side and at the later gas passes, using cylindrical elements along the longitudinal direction. A schematic representation of the FVM modeling of the flue gas side is provided in Fig. 1.

The number of finite volume elements is selected considering the non-linearities introduced by the time-based differential algebraic equations, while balancing the computational cost. Accordingly, 100 volume elements N_{CC} have been considered for the combustion chamber, while $N_{2nd} = N_{3rd} = 20$ elements are utilized for the 2nd and 3rd gas passes. For solving the equations, a numerical solver for stiff problems of the ODEs, originally implemented by Shampine et al. [49] has been employed.

Combustion chamber

For each *i*th volume element, the main dynamic heat balance expression describes the resulting spatially homogeneous temperature (of

the discrete flue-gas cylindrical volume), as the net result of the energy terms involved in the combustion process and the heat exchange with the surrounding metal wall, which is expressed in Eq. (1).

$$\frac{d}{dt}(\rho_g V c_{V,g,CC} T^i_{g,CC}) = \dot{m}_{fuel} LHV \Delta Y_i + \dot{m}_g \bar{c}_{P,g,CC}(T^{i-1}_{g,CC} - T^i_{g,CC}) - \dot{q}^i_{rad,CC} - \dot{q}^i_{conv,CC} \quad (1)$$

Where the wall temperature is assumed to be constant, which is motivated by the predominant heat conduction effect of the metal surface [50]. The discretized fractional heat release rate by combustion ΔY, is instead expressed in Eq. (2).

$$\Delta Y_i = exp\left\{-4.6\frac{i}{N_{CC}}\right\}\left(exp\left\{\frac{4.6}{N_{CC}}\right\} - 1\right) \quad (2)$$

Furthermore, the convective heat transfer flow term is reported in Eq. (3):

$$\dot{q}^i_{conv,CC} = h^i_{conv,CC}\frac{2\pi R_{CC} L_{CC}}{N_{CC}}(T^i_{g,CC} - T_{wall,CC}) \quad (3)$$

Where the convective heat transfer coefficient is estimated employing the Dittus–Boelter equation for internal cooling flow of flue-gases [23], while the flow's properties are described assuming an ideal gas mixture behavior. The convective heat transfer coefficient is therefore expressed analytically in Eq. (4).

$$h^i_{conv,CC} = \frac{0.023 \cdot k_{g,CC}}{2 \cdot R_{CC}} Re_i^{0.8} Pr_i^{0.3} \quad (4)$$

Next, the radiative heat flux, imposed by the longitudinal shaped flame, is modeled assuming the metal walls to be gray Lambertian bodies, while the extracted flame's fuel dependant emissivity is modeled thanks to the Hottel and Sarofim [22] correlation. For the current application, natural gas isotropic emissivity is assumed to be equal to 0.85, where the net radiative heat flux is reported in Eq. (5).

$$\dot{q}^i_{rad,CC} = \frac{2\pi R_{CC} L_{CC}}{N_{CC}}\sigma\frac{\varepsilon_{wall,CC} + 1}{2}\varepsilon_{g,CC}(T^{4,i}_{g,CC} - T^4_{wall,CC}) \quad (5)$$

Further details and a demonstration about the share of the radiative heat flux, along with the dynamic energy balance applied to the combustion chamber and the details of the model employed for simulating the 2nd and 3rd gas passes, are presented in Appendix D (provided in supporting materials).

Water side

Alongside the net heat quantity transferred through conduction by the metal separation walls, two additional energy flow terms (that cross the boundary of the boiler vessel) are considered. These terms correspond to the sub-cooled water's inlet and outlet flows that dynamically interact with the mentioned water control volume. Liquid water is assumed to be incompressible; therefore, no dynamic variation of the water volume inside the reservoir is considered. The volumetric feedwater flow rate is externally controlled by a centrifugal pump that is managed by an external control entity, which is outside the scope of the present investigation. The boiler vessel is geometrically designed to promote the mixing phenomena of the water flows and the local heat fluxes of the previously mentioned heat transfer regions. Furthermore, the boiler's temperature is assumed to be homogeneously distributed. The equation employed for modeling the heat transfer between the tube-bank and the mixing water, along with additional details about the water side's model, are presented in Appendix D (provided in supporting materials). The approach utilized to determine the boiler's efficiency is also reported in Appendix D (provided in supporting materials).

Data-driven modeling

As was previously pointed out, in many industrial case studies (specifically those in which old components are utilized or those in which the simulation is being conducted by a third party), the geometrical details of the boiler are not available; a data-driven approach, utilizing the data obtained from the boiler, can thus be utilized to simulate its dynamic behavior. Accordingly, in the present study, in order to accurately forecast the boiler's thermal behavior, an ML [51] based pipeline aimed to estimate the outlet temperature of the hot-water is developed. The pipeline is developed in the training step and is then assessed in the testing phase. Accordingly, the available processed dataset is first divided into the training and testing sub-sets. In the training step, in which only the training sub-set is utilized, the training algorithm of the machine-learning model, iteratively adapts its internal configuration, subject to the minimization of the target error function with respect to the experimental training data. The latter procedure is conducted through performing the cross-validation (CV) algorithm [52,53], in which the training dataset itself is divided into k equal subsets. The model is next trained utilizing $k-1$ subsets and is then validated by predicting the target data with the remaining k subset [54]. By repeating the mentioned process for k iterations, all the samples of the training data set are utilized as train and validation points at least once. Employing the latter approach, the models are less prone to over-fitting [55] issue. In the testing step, the rest of the input dataset is employed to compute the performance of the resulting trained data-driven model in predicting the desired target [51]. The initial steps of the implemented approach include performing the data cleaning and filtering followed by the creation of lagged features. In this context, since the developed models should simulate the boiler while it is in operation (On state), the available data is filtered preserving only the recordings in which the natural gas flow rate is more than zero. Next, pipelines with different prediction horizons are implemented and the performance of three benchmark algorithms (Random Forest (RF) [41], Extreme Gradient Boosting (XGBoost), and Support Vector Machines (SVM)), for each pipeline is determined. Subsequently, an algorithm optimization step is performed identifying the most promising algorithm (and the corresponding tuning parameters) for each pipeline. Finally, the performance of the initial (with benchmark algorithms) and optimized pipelines, utilizing the considered accuracy metrics, are compared. The employed accuracy metrics along with brief descriptions of utilized algorithms are provided in Appendix E (provided in supporting materials).

Control strategy

As was indicated in the description of the hot-water boiler and the corresponding dynamic model, the system has two degree of freedom: the fuel-air mixture flow rate and the water flow rate. The system's dynamics is controlled through the management of the these variables, which thus determines (described by the provided dynamic equations) the temperature of the water inside the boiler vessel. The feedwater flow rate is traditionally controlled by circulating centrifugal pumps, which operate at a rotating velocity that is driven by the thermal demand (downstream of the boiler). Therefore, the water flow control apparatus is separate from the boiler control's boundaries. The combustion rate is instead actively controlled by an agent, the control variable of which is the water temperature inside the boiler. The mentioned agent's architecture is traditionally composed of a PID controller. Moreover, the closed-loop control estimates the bias between the water set-point temperature and the mentioned process variable.

Following the described control architecture, the PID controller's parameters need to be tuned in order to minimize the undesired hysteresis that would in turn reduce the hot-water supply quality (deviation from the desired setpoint) and the system's efficiency, which is

motivated by the rapid variation of fuel flow rate that increases the flue-gas's discharge temperature. Among the available tuning procedures, a trial and error process is commonly employed utilizing time-consuming and resource-intensive experiments. The mentioned methodology is labour intensive and requires the presence of a testing apparatus. A different methodology, which has been implemented in the present work (as described at Section "Mathematical model"), is conducted through the linearization procedure. Plant's local linearization is made available thanks to the presence of detailed dynamic model of the system. The linearization operation defines the plant's transfer function employing the estimated input and output variables [56]. The procedure thus allows a drastic reduction of the (otherwise necessary) fuel and labour costs, while resulting in an improvement compared with the sub-optimal PID coefficients that are obtained in the mentioned experimental procedure. Furthermore, the obtained optimal controller's coefficients cover a wide range of boiler's operating conditions in contrast to the sub-optimal solution, which take into account only the boiler's nominal performances. Thus, an improved control stability, outside the maximum power output state, is also achieved.

Results and discussions

In the present section, the results of the physical model's validation and corresponding considerations are first reported and discussed. In the second step, the data-driven predictive modeling approach is implemented and the achieved performance of the chosen benchmark ML algorithms, for the implemented pipelines (with different prediction horizons), are presented. Finally, the optimal pipeline, for each prediction horizon, is determined and the resulting performance (estimation error) is compared with those offered by the above-mentioned benchmark algorithms.

Physical model's validation

In order to validate the developed physical dynamic model, the measurements performed on a hot-water boiler that is utilized in an Italian industrial factory (including the hot-water consumption data and boiler's operating parameters) have been used. The considered boiler is employed in the factory for space-heating (in the winter) and warming up the raw materials (later utilized in manufacturing process). In this context, the measurement data corresponding to the month of January has been utilized, as this period (due to the elevated thermal demand for space heating in northern Italy) includes the highest required thermal production, which (as illustrated at Fig. 2 (right)) helps evading the long-term stand-by events of the boiler. The water volumetric flow rate, temperature of the feedwater injected into the boiler vessel, along with the supplied water's temperature are measured, which are correspondingly reported at Fig. 2 (left). At the flue-gas side, the corrected natural gas flow rate is registered at the standard conditions, while a gas analyzer measures the volumetric O_2 and CO concentration at the stack alongside the flue-gas discharge temperature. The measured variables (as listed in Table 1) allow performing an energy analysis on the system as a whole and thus permit the estimation of the boiler's efficiency. The sampling frequency is set equal to 1 min, allowing fine granularity of the measurement, thus retaining major information about the system's dynamic behavior.

The external operating conditions that are imposed, while performing the validation process, are reported in Table 1. Additionally, the registered average stack oxygen concentration is directly employed while simulating the combustion process to coherently estimate (in a dynamic manner) the combustion quality. The additional geometrical specifications, of the considered boiler model, are listed in Table 1. For the previously described consumption profile of hot-water flow rate, the performance of the boiler is estimated and the corresponding key indicators are obtained. It is also worth mentioning that the model's sub-parts validation has been carried out by the authors in previous works [3,19,36], therefore the current procedure aims at providing the boiler model's validity as a whole, thus considering the additional water-side and control strategy parts in addition to the previously presented works. The cumulative average efficiency is a performance index, which has also been utilized as the primary validation term (compared with the experimental data). Next, in order to evaluate the deviation from the registered variables of the case-study, the estimation bias of the water supply temperature and the mass (cumulative) of consumed fuel (determined by comparing the measured data and the estimated results) is calculated.

Table 1
List of experimental variables, geometrical characteristics, Ambient conditions and the range of boiler's operating conditions and their instrument errors recorded on the hot-water boiler.

Variables recorded	Instrument error
Water hourly volumetric flow rate m^3/h	±1% o.r.
Inlet water feed temperature °C	±0.4 °C
Outlet water feed temperature °C	±0.4 °C
Natural gas volumetric hourly flow rate Sm^3/h	±1% o.r.
Flue gas stack temperature °C	±0.4 °C
Stack percentage volumetric average oxygen concentration $V_{O_2}/V_{tot,f}$%	±1% o.r.
Stack percentage volumetric average CO concentration ppm	±10 ppm
Parameters of the simulation	**Values**
Ambient temperature [°C]	20
Ambient pressure [bar]	1.01325
Ambient relative humidity [%]	50.0
Maximum hourly fuel mass flow rate [kg/h]	100.0
Parameters of the Fire-tube boiler	**Values**
Number of tubes (3rd gas pass)	116
Helical coils pitch (3rd gas pass) [mm]	80
Helical coils diameter (3rd gas pass) [mm]	42
Boiler net vessel volume [m^3]	3.65
Heat transfer surface [m^2]	86.10
Net water volume/Heat transfer surface ratio [10^3 m]	42.39

Table 2
The bias between simulated and experimentally measured (and calculated) thermal efficiency and (cumulative) consumed fuel mass and the bias indicators obtained comparing the experimental and simulated data.

	Experimental	Simulated
Cumulative thermal efficiency [%]	104.21	103.01
Mass (cumulative) of consumed fuel [kg]	186.13	191.82
	Cumulative bias	
Water supply temperature [K]	0.4126	
Natural gas flow rate [kg/h]	6.2192	

Validation results

Table 2 compares the simulated cumulative thermal efficiency and the (cumulative) mass of consumed fuel with the corresponding experimental values. Results demonstrate an acceptable agreement between the simulated values and the measured data, indicating a cumulative thermal efficiency deviation of 1.2%.

Fig. 3 represents the trends of simulated natural gas flow rate and the supply water temperature in comparison with the corresponding experimental data. The latter comparison demonstrates the marginal deviation of the simulated data with respect to the experimental measurements, confirming the model's validity over the range of operating conditions that has been investigated in the present work. Furthermore, the resulting cumulative deviations (simulation results vs. the measured value) of these two parameters is reported in Table 2, which further confirm the consistency of the simulation results with the experimental data.

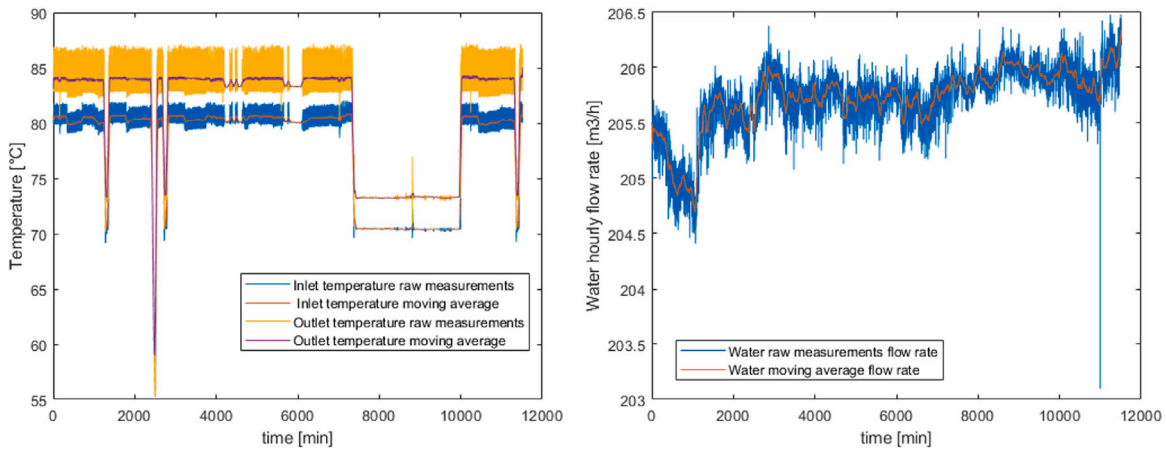


Fig. 2. (Left) Water volumetric flow rate trend within one week in January. (Right) Water inlet and outlet temperature trends experimentally recorded within one week in January.

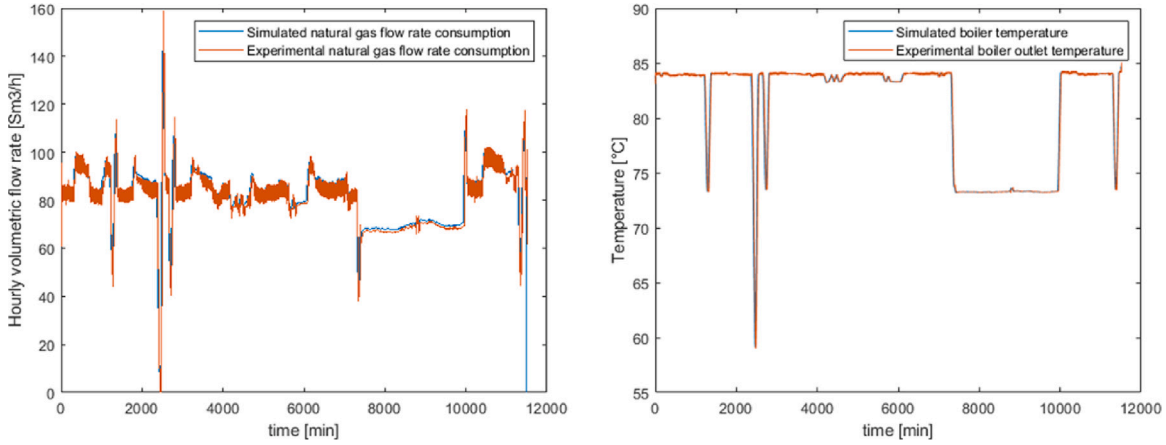


Fig. 3. Comparison between the simulated natural gas flow rate and experimental data (left) and simulated water supply temperature vs. the experimental data (right).

Results of data-driven dynamic modeling

In the present section, the results of data-driven predictive modeling procedure, which aims at forecasting the water supply temperature with different prediction horizons (15 min, 30 min, and one hour), are represented. The experimental data, described in sub-section "Physical model's validation", is utilized for training, validating, and testing the implemented prediction pipelines. In the data creation step (to describe the underlying dynamic physical phenomena), lagged values of the target and input features, for each pipeline (while taking into account the corresponding prediction horizon), are created and are added to the dataset. The dataset is then split into training (also used for validation through cross-validation) and test subsets. This splitting step is performed by choosing the training set (that constitutes 70% of the dataset) and the test set (that constitutes 30% of the dataset) to be (chronologically) subsequent.

Next, utilizing time-series cross-validation methodology (with 10 chronologically subsequent splits), the performance of chosen benchmark ML algorithms (SVM, XGBoost, and RF), for each prediction pipeline, are assessed (validation accuracy). Next, the accuracy of the latter pipelines for the test set (while training the model over the whole training dataset) is determined. The achieved validation accuracy of these models for different prediction horizons (in terms of MAE, MSE, MARD, and R^2 metrics) are represented in Table 3. The corresponding results for the test set are represented in Table 3. In the next step, for each pipeline, the algorithm optimization procedure, which is described in Appendix E (provided in supporting materials), is performed. In this procedure, while utilizing MARD as the optimization objective and the training (also used for validation through cross-validation) subset as the dataset, the most promising ML-based pipeline for each prediction horizon is identified. Finally, using the identified optimal pipelines, the corresponding accuracy over the test set is calculated. The motivation behind the choice of MARD as the key objective in the optimization procedure is the fact that (in the considered application) it is desired to reduce the percentage of the difference between the real and predicted values (while normalizing the impact of the magnitude of the real value). It is also worth mentioning that splitting the dataset into training and test sets, facilitates assessing the performance of the identified optimal pipeline for a (unseen) dataset for which it has not been optimized.

Comparing the validation MARD value achieved by the optimal pipeline with those obtained utilizing the benchmark algorithms (represented in Table 3), demonstrates that conducting the pipeline optimization step, results in a notable reduction in the resulting estimation error. In addition, comparing the MARD results of benchmark algorithms for the validation (cross-validation over the training set) and test sets demonstrates that the overall estimation error is higher for the test set independent from the choice of the model. The latter observation can be attributed to the difference in the profile of the demand in the test set that, due to the resulting increase in temperature oscillations, makes the supply temperature prediction task more challenging. Accordingly, the achieved MARD values of the optimal pipeline are similarly higher for the test set compared to the corresponding achieved values for the validation set. Therefore, this difference can be partly attributed to the above-mentioned difference in the supply temperature profile in the two dataset and to a lower extent to the potential model over-fitting.

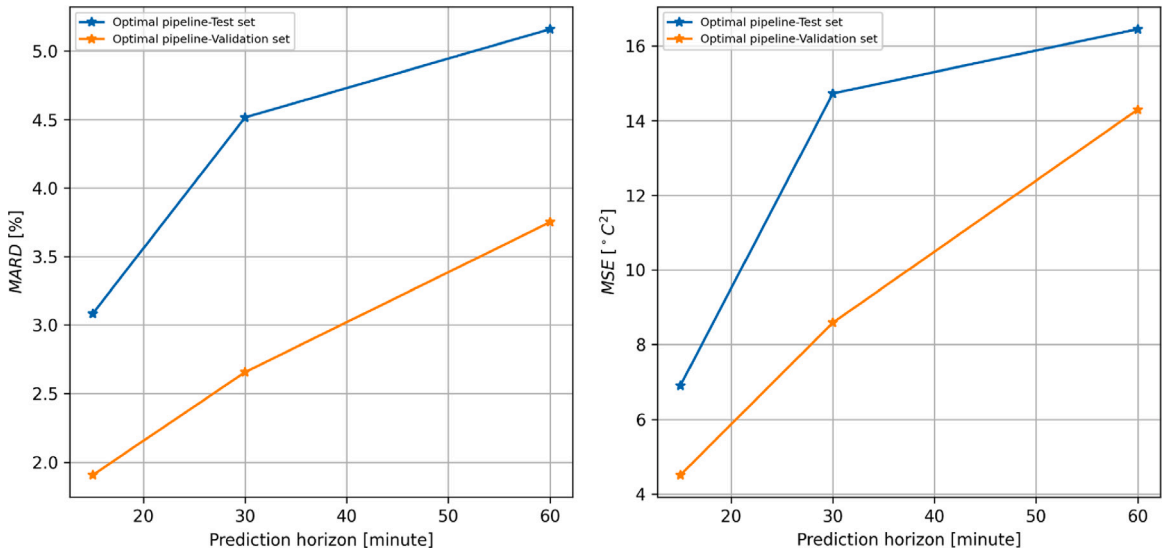


Fig. 4. Effect of various prediction horizons on evaluation metrics of the validation and the test set.

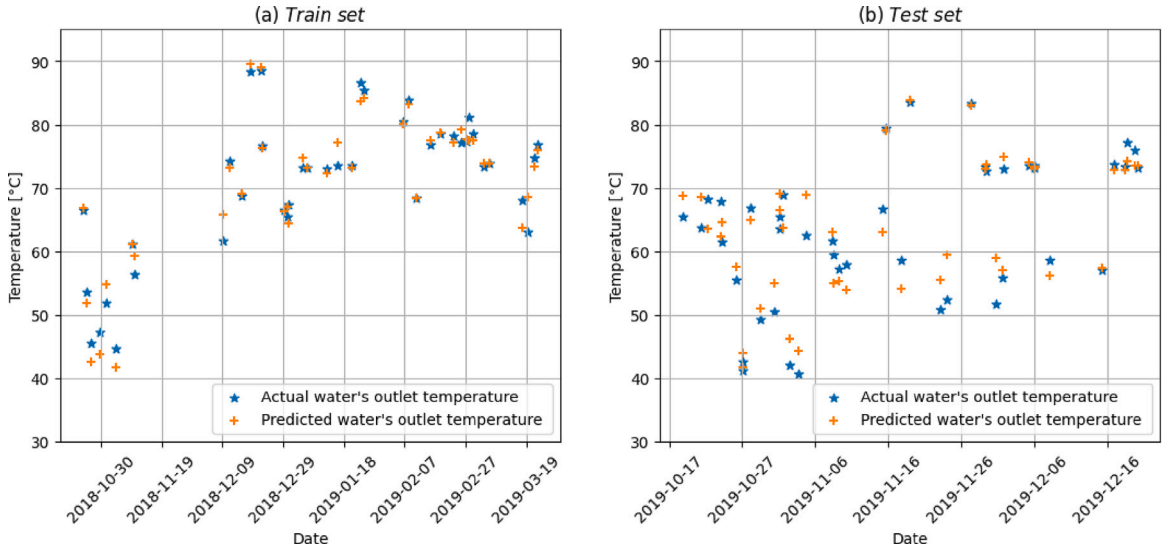


Fig. 5. Accuracy of the 30-minute optimal pipeline predictive model on predicting train and test dataset.

Nevertheless, the achieved MARD of the optimal pipeline on the test set is notably lower than those offered by the state-of-the-art benchmark algorithms for all of the considered prediction horizons, which demonstrates the notable benefit of implementing the pipeline optimization procedure. Fig. 4 graphically represents the above-mentioned observations on the variations of the estimation error achieved by the optimal pipeline over the validation and test sets for different prediction horizons.

In order to visually represent the accuracy of the achieved optimal pipelines, the comparison between the estimations provided by the identified optimal pipeline for 30 min ahead prediction and the corresponding real values is illustrated in Fig. 5. As can be noticed in this figure, the estimations of the optimal pipeline are in a suitable agreement with the experimentally measured data. The optimal pipeline can estimate the values of the outlet water temperature in the next 30 min with an MAE of less than 1.9 °C for the validation and 2.7 °C for the test set, which demonstrates its elevated forecasting accuracy. These pipelines can thus be effectively utilized in a model predictive control scheme, aiming at improving the controllability of the boiler and potentially reducing its consumption. In order to provide the reader with complete information, the details of the identified obtained optimal pipelines (including various steps, algorithms, and their tuned hyper parameters along with the corresponding definitions), for different prediction horizons, are reported in Appendix A to Appendix C (provided in supporting materials).

Conclusions

The present work was aimed at providing dynamic thermal models of hot-water supplying fire-tube boilers. As the first step, a detailed dynamic physical model, employing reduced one-dimensional finite-volume methodology to simulate the flue-gas side, was proposed and implemented. In the next step, the measurements performed on the demand profile and the operating conditions of a hot-water boiler installed in an Italian industrial factory, were utilized for both tuning a PID controller (using a closed-loop linearization procedure) and the validation of the developed model. The obtained results demonstrated a maximum thermal efficiency bias (determined through comparing the experimental data and the simulated values) of 1.2%, which confirms the validity of the developed model over the investigated range of

Table 3
Results of Cross-Validation analysis and test results applying SVM, XGBoost, RF, and the optimal pipeline models of different prediction horizons.

Prediction horizon	Evaluation metric	SVM	XGBoost	RF	Optimal pipelines
		Cross-validation results			
15 min	MARD [%]	2.41	3.10	3.39	1.90
	MAE [°C]	1.4	1.9	2.1	1.2
	MSE [°C^2]	4.8	6.9	10.6	4.5
	R^2 [−]	0.92	0.95	0.93	0.88
30 min	MARD [%]	3.55	5.44	5.01	2.66
	MAE [°C]	2.9	3.3	3.1	1.9
	MSE [°C^2]	11.7	22.2	18.4	8.6
	R^2 [−]	0.83	0.84	0.87	0.80
60 min	MARD [%]	6.40	7.86	6.62	3.75
	MAE [°C]	4.4	5.1	4.2	2.9
	MSE [°C^2]	24.5	53.8	37.5	14.3
	R^2 [−]	0.75	0.62	0.74	0.73
		Test results			
15 min	MARD [%]	3.90	5.89	5.94	3.08
	MAE [°C]	2.8	2.9	3.1	1.9
	MSE [°C^2]	7.2	30.1	30.2	6.8
	R^2 [−]	0.96	0.61	0.55	0.95
30 min	MARD [%]	6.10	7.53	6.13	4.51
	MAE [°C]	3.2	3.3	3.0	2.7
	MSE [°C^2]	17.3	49.2	29.0	14.7
	R^2 [−]	0.88	0.55	0.57	0.84
60 min	MARD [%]	8.70	13.44	9.11	5.16
	MAE [°C]	5.9	6.6	5.3	3.4
	MSE [°C^2]	26.4	184.9	91.5	16.5
	R^2 [−]	0.87	0.46	0.48	0.76

operating conditions. The model can thus be effectively utilized in the design and sizing procedure of this type of boilers.

In the second part of the study, ML-based pipelines were implemented and optimized for predicting the water supply temperature with 15 min, 60 min, and one hour prediction horizons. It was demonstrated that performing the pipeline optimization step can notably reduce the estimation error of the prediction pipeline (compared with the state-of-the-art benchmark ML algorithms) for all of the considered prediction horizons. It was specifically shown that the MARD values (on the test) achieved by identified optimal pipelines are 3.1%, 4.5%, and 5.2% for 15 min, 30 min, and one hour prediction horizons respectively, which demonstrates their elevated accuracy. Therefore, these pipelines can be utilized for implementing predictive control schemes to improve the controllability and (potentially) the thermal efficiency of these boilers.

CRediT authorship contribution statement

Marco Tognoli: Methodology, Software, Investigation, Formal analysis. **Shayan Keyvanmajd:** Writing – original draft, Visualization, Validation. **Behzad Najafi:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Fabio Rinaldi:** Project administration, Resources, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

The authors would like to thank ICI Caldaie S.p.A. for providing technical and financial support for this project.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.seta.2023.103321.

References

[1] Harris K. Model boilers and boilermaking. MAP technical publication, Model and Allied Publications; 1974.
[2] Tognoli M, Najafi B, Rinaldi F. Dynamic modelling and optimal sizing of industrial fire-tube boilers for various demand profiles. Appl Therm Eng 2018;132:341–51.
[3] Tognoli M, Najafi B, Marchesi R, Rinaldi F. Dynamic modelling, experimental validation, and thermo-economic analysis of industrial fire-tube boilers with stagnation point reverse flow combustor. Appl Therm Eng 2019;149:1394–407.
[4] Colonna vP, Van Putten H. Dynamic modeling of steam power cycles.: Part I—Modeling paradigm and validation. Appl Therm Eng 2007;27(2–3):467–80.
[5] De Mello F. Boiler models for system dynamic performance studies. IEEE Trans Power Syst 1991;6(1):66–74.
[6] Sun L, Li D, Lee KY, Xue Y. Control-oriented modeling and analysis of direct energy balance in coal-fired boiler-turbine unit. Control Eng Pract 2016;55:38–55.
[7] Adam E, Marchetti J. Dynamic simulation of large boilers with natural recirculation. Comput Chem Eng 1999;23(8):1031–40.
[8] Åström KJ, Bell RD. Drum-boiler dynamics. Automatica 2000;36(3):363–78.
[9] Kim H, Choi S. A model on water level dynamics in natural circulation drum-type boilers. Int Commun Heat Mass Transfer 2005;32(6):786–96.
[10] Blanco JM, Vazquez L, Peña F. Investigation on a new methodology for thermal power plant assessment through live diagnosis monitoring of selected process parameters; application to a case study. Energy 2012;42(1):170–80.
[11] Chandrasekharan S, Panda RC, Natrajan Swaminathan B. Dynamic analysis of the boiler drum of a coal-fired thermal power plant. Simulation 2017;93(11):995–1010.
[12] Zhou Y, Wang D. An improved coordinated control technology for coal-fired boiler-turbine plant based on flexible steam extraction system. Appl Therm Eng 2017;125:1047–60.
[13] El-Guindy A, Rünzi S, Michels K. Optimizing drum-boiler water level control performance: A practical approach. In: 2014 IEEE conference on control applications. CCA, IEEE; 2014, p. 1675–80.
[14] Oko E, Wang M, Zhang J. Neural network approach for predicting drum pressure and level in coal-fired subcritical power plant. Fuel 2015;151:139–45.
[15] Coelho PJ, Novo PA, Carvalho M. Modelling of a utility boiler using parallel computing. J Supercomput 1999;13(2):211–32.
[16] Westbrook CK, Mizobuchi Y, Poinsot TJ, Smith PJ, Warnatz J. Computational combustion. Proc Combust Inst 2005;30(1):125–57.
[17] Gómez A, Fueyo N, Diez LI. Modelling and simulation of fluid flow and heat transfer in the convective zone of a power-generation boiler. Appl Therm Eng 2008;28(5–6):532–46.
[18] Pezo M, Stevanovic VD, Stevanovic Z. A two-dimensional model of the kettle reboiler shell side thermal-hydraulics. Int J Heat Mass Transfer 2006;49(7–8):1214–24.
[19] Morelli A, Tognoli M, Ghidoni A, Najafi B, Rinaldi F. Reduced FV modelling based on CFD database and experimental validation for the thermo-fluid dynamic simulation of flue gases in horizontal fire-tubes. Int J Heat Mass Transfer 2022;194:123033. http://dx.doi.org/10.1016/j.ijheatmasstransfer.2022.123033.
[20] Sørensen K, Karstensen CM, Condra T, Houbak N. Modelling and simulating fire tube boiler performance. In: Proceedings of SIMS. 2003, p. 9–12.
[21] Ortiz FG. Modeling of fire-tube boilers. Appl Therm Eng 2011;31(16):3463–78.
[22] Hottel H, Sarofim A. Radiative transfer, New York. 1967, p. 20–4.
[23] Dittus F, Boelter L. Heat transfer in automobile radiators of the tubular type. Int Commun Heat Mass Transfer 1985;12(1):3–22. http://dx.doi.org/10.1016/0735-1933(85)90003-X, URL http://www.sciencedirect.com/science/article/pii/073519338590003X.
[24] Cooper M. Heat flow rates in saturated nucleate pool boiling-a wide-ranging examination using reduced properties. In: Advances in heat transfer, vol. 16. Elsevier; 1984, p. 157–239.
[25] Smith T, Shen Z, Friedman J. Evaluation of coefficients for the weighted sum of gray gases model. 1982.
[26] Abene A, Rahmani A, Seddiki RG, Moroncini A, Guillaume R. Heat transfer study in 3-pass fire-tube boiler during a cold start-up. Softw Eng 2018;5(4):57.
[27] Churchill SW, Chu HH. Correlating equations for laminar and turbulent free convection from a horizontal cylinder. Int J Heat Mass Transfer 1975;18(9):1049–53.
[28] Ji W-T, Jacobi AM, He Y-L, Tao W-Q. Summary and evaluation on the heat transfer enhancement techniques of gas laminar and turbulent pipe flow. Int J Heat Mass Transfer 2017;111:467–83.
[29] Sethumadhavan R, Rao MR. Turbulent flow heat transfer and fluid friction in helical-wire-coil-inserted tubes. Int J Heat Mass Transfer 1983;26(12):1833–45.
[30] Klaczak A. Heat transfer in tubes with spiral and helical turbulators. 1973.

[31] Inaba H. A fundamental study of heat-transfer enhancement and flow-drag reduction in tubes by means of wire coil insert: 1st report, characteristics of flow resistance and heat transfer in tubes wiht wire coil insert. JSME B 1994;60:240–7.
[32] Uttarwar S, Raja Rao M. Augmentation of laminar flow heat transfer in tubes by means of wire coil inserts. 1985.
[33] Garcia A, Vicente PG, Viedma A. Experimental study of heat transfer enhancement with wire coil inserts in laminar-transition-turbulent regimes at different Prandtl numbers. Int J Heat Mass Transfer 2005;48(21–22):4640–51.
[34] Kasturi ML, Junaid M, Awate YS, Acharya AR. Effect of wire coil turbulators on pressure drop and heat transfer augmentation in a circular tube. In: 2017 International conference on innovations in information, embedded and communication systems. 2017, p. 1–6. http://dx.doi.org/10.1109/ICIIECS.2017.8276115.
[35] Sharafeldeen M, Berbish N, Moawed M, Ali R. Experimental investigation of heat transfer and pressure drop of turbulent flow inside tube with inserted helical coils. Heat Mass Transf 2017;53(4):1265–76.
[36] Morelli A, Tognoli M, Ghidoni A, Najafi B, Rinaldi F. Reduced thermal-hydraulic model of flue gases in horizontal fire tubes with helical coil inserts employing a CFD simulated database. SSRN 2022. URL https://ssrn.com/abstract=4210318.
[37] Tognoli M, Najafi B, Lucchini A, Colombo LPM, Rinaldi F. Implementation of a multi-setpoint strategy for fire-tube boilers utilized in food and beverage industry: Estimating the fuel saving potential. Sustain Energy Technol Assess 2022;53:102481. http://dx.doi.org/10.1016/j.seta.2022.102481.
[38] Dadras Javan F, Khatam Bolouri Sangjoeei H, Najafi B, Haghighat Mamaghani A, Rinaldi F. Application of machine learning in occupant and indoor environment behavior modeling: Sensors, methods, and algorithms. Handbook of smart energy systems. Springer; 2022, p. 1–25. http://dx.doi.org/10.1016/B978-0-12-814761-0.00003-4.
[39] Ardam K, Najafi B, Lucchini A, Rinaldi F, Colombo LPM. Machine learning based pressure drop estimation of evaporating R134a flow in micro-fin tubes: Investigation of the optimal dimensionless feature set. Int J Refrig 2021;131:20–32.
[40] Najafi B, Ardam K, Hanušovský A, Rinaldi F, Colombo LPM. Machine learning based models for pressure drop estimation of two-phase adiabatic air-water flow in micro-finned tubes: Determination of the most promising dimensionless feature set. Chem Eng Res Des 2021;167:252–67.
[41] Manivannan M, Najafi B, Rinaldi F. Machine learning-based short-term prediction of air-conditioning load through smart meter analytics. Energies 2017;10(11).
[42] Romeo LM, Gareta R. Neural network for evaluating boiler behaviour. Appl Therm Eng 2006;26(14–15):1530–6.
[43] Smrekar J, Pandit D, Fast M, Assadi M, De S. Prediction of power output of a coal-fired power plant by artificial neural network. Neural Comput Appl 2010;19(5):725–40.
[44] Selleri T, Najafi B, Rinaldi F, Colombo G. Mathematical modeling and multi-objective optimization of a mini-channel heat exchanger via genetic algorithm. J Thermal Sci Eng Appl 2013;5(3).
[45] Najafi B, Najafi H, Idalik M. Computational fluid dynamics investigation and multi-objective optimization of an engine air-cooling system using genetic algorithm. Proc Inst Mech Eng C 2011;225(6):1389–98.
[46] Najafi B, Di Narzo L, Rinaldi F, Arghandeh R. Machine learning based disaggregation of airconditioning loads using smart meter data. IET Gener Transm Distrib 2020;14(21):4755–62.
[47] Annaratone D. Steam generators: description and design. 2008, http://dx.doi.org/10.1007/978-3-540-77715-1.
[48] Rinaldi F, Najafi B. Temperature measurement in WTE boilers using suction pyrometers. Sensors (Switzerland) 2013;13(11):15633–55.
[49] Shampine LF, Reichelt MW. The matlab ode suite. SIAM J Sci Comput 1997;18(1):1–22.
[50] Flynn M, O'Malley M. A drum boiler model for long term power system dynamic simulation. IEEE Trans Power Syst 1999;14(1):209–17.
[51] Najafi B, Moaveninejad S, Rinaldi F. Data analytics for energy disaggregation: methods and applications. In: Big data application in power systems. Elsevier; 2018, p. 377–408.
[52] Stone M. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Ser B Stat Methodol 1974;36(2):111–33.
[53] Campodonico Avendano IA, Dadras Javan F, Najafi B, Moazami A, Rinaldi F. Assessing the impact of employing machine learning-based baseline load prediction pipelines with sliding-window training scheme on offered flexibility estimation for different building categories. Energy Build 2023;294:113217. http://dx.doi.org/10.1016/j.enbuild.2023.113217.
[54] Najafi B, Depalo M, Rinaldi F, Arghandeh R. Building characterization through smart meter data analytics: Determination of the most influential temporal and importance-in-prediction based features. Energy Build 2021;234:110671. http://dx.doi.org/10.1016/j.enbuild.2020.110671.
[55] Anderson D, Burnham K. Model selection and multi-model inference, vol. 63, no. 2020. Second NY: Springer-Verlag; 2004, p. 10.
[56] Groves K, Serrani A. Modeling and nonlinear control of a single-link flexible joint manipulator. Appendices 2004.