Knowledge-Based Systems 000 (2016) 1–11

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A globally enhanced general regression neural network for on-line multiple emissions prediction of utility boiler

Jingge Song [a], Carlos E. Romero [b], Zheng Yao [b,*], Boshu He [a]

[a] *Institute of Combustion and Thermal Systems, School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, No.3 Shangyuancun, Beijing 100044, China*
[b] *Energy Research Center, Lehigh University, 117 ATLSS Drive, Bethlehem, PA 18015, USA*

ARTICLE INFO

Article history:
Received 22 March 2016
Revised 1 November 2016
Accepted 4 November 2016
Available online xxx

Keywords:
General regression neural network
Gaussian adapted resonance theory
Wiggle-method
Extrapolation
On-line combustion modeling
Multiple emissions prediction

ABSTRACT

Prediction of multiple emissions of boiler such as NO_X, SO_2 and mercury accurately and quickly can help operating engineers understanding the power unit deeply and controlling harmful emissions efficiently. However, most of existing works did not consider multiple emissions prediction, on-line prediction and accurately prediction in a simple and uniform framework. To this end, a new neural network system, named Globally Enhanced General Regression Neural Network (GE-GRNN), is proposed to solve multiple emissions prediction problem for utility boiler under on-line environment. The proposed GE-GRNN is based on General Regression Neural Network (GRNN), and employs Gaussian Adapted Resonance Theory (GART) as an incremental learning method to reduce memory cost of GRNN emission model, which is very suitable for large-scale real-time input samples. As contributions, three methods are introduced in GE-GRNN: (1) a modified wiggle-method is proposed to adjust smooth factors of GRNN dynamically to enhance both global and local estimations; (2) a fast polynomial extrapolation structure is designed in hidden layers of GRNN to improve the quality of extreme value estimation; and (3) a hybrid estimation mechanism is established to integrate wiggle-method and extrapolation into an uniform estimation framework. The simulation experiments are conducted on mathematical functions and real-world emissions predication of NO_X and loss on ignition (LOI) of fly ash in a 600MW boiler. Results show that the proposed system performs attractive on-line performance while keeping agreement with testing samples well. The predicted emission performance of the tested boiler is reasonable, and can provide valuable reference to emission optimization. The proposed framework is domain independent and can be used to other fields for on-line multiple performances prediction.

1. Introduction

With extensively attentions been paid to environmental protection worldwide, emissions control in power plant is becoming an important and urgent requirement during boiler operation [1]. Emission prediction, an important technology for both improving boiler combustion and the operation of emission control equipment such as selective catalytic reduction (SCR) and flue gas desulfurization (FGD), has received continuous interests from researchers [1–5]. In real process, boiler combustion is a complex non-linear process that can produce multiple harmful emissions such as NO_X, CO, SO_2 and mercury, and is also influenced by many factors such as fuel quality, air supply, burner type and over-fired air (OFA). The prediction of multiple emissions in on-line environment is a challenge work and can provide more comprehensive understanding of the behavior of boiler combustion. The purpose of this study is to investigate an efficient knowledge-based framework for on-line prediction of multiple emissions accurately, quickly and with self-learning ability.

Data-driven intelligent methods, such as Artificial Neural Network (ANN) and Supported Vector Machine (SVM), have been accepted as powerful tools for modeling and predicting complex science and engineering process. Boiler combustion, a major source of the air pollution and the resultant greenhouse gas (GHG) emissions, is influenced by many factors such as fuel quality, air supply, burner type, burner tilt, over-fired air (OFA) opening, measurement instrumentation, control logic and the operator's experience. In the late decades, numerous researches have been conducted with regard to prediction the performance of boiler emissions based on ANN or SVM. Zhou et al. [1] developed an optimized Backward Propagation Neural Network (BPNN) model for NO_X emission prediction and control of a 600MW utility boiler. Chu et al. [3] predicted NO_X and CO emissions with the help of a three layers

* Corresponding author.
E-mail address: zheng.yao@lehigh.edu (Z. Yao).

http://dx.doi.org/10.1016/j.knosys.2016.11.003
0950-7051/

common ANN and simulated dataset. Abdel-Aal [2] predicted mercury speciation in utility boiler flue gases using GMDH-based abductive networks and compared its performance with ANN. Ilamathi et al. [6] used BPNN for estimating unburned carbon in bottom ash of a tangentially fired 210 MW boiler. Krzywanski et al. [7] developed a compact ANN model for prediction of SO_2 emission in a 261MW Circulating Fluidized Bed (CFB) boiler with a lot of input parameters. On the other hand, Si et al. [8] developed a modified structural adaptive on-line SVM (AOSVR) for NO_X emission control, Gu et al. [9] integrated several improved least squares SVMs (LS-SVM) for boiler combustion modeling and optimization, and Zhou et al. [10] also used SVM for boiler NO_X emission modeling. ANN and SVM have their own advantages and disadvantages. Theoretically, a three-layer feed forward ANN can approximate any nonlinear function; establish potential relationship between the several output parameters such as multiple emissions with only one network; however, is limited on generalization ability. SVM has good generalization ability, but for large-scale samples the training is time consuming. Moreover, a SVM can only estimate one output. For multiple outputs estimation, multiple SVMs need to be constructed separately, thus potential relationship among outputs might be ignored [9]. In recent years, Radial Basis Function Neural Network (RBF-NN) received more attentions for the advantages such as avoiding falling into local optimum easily and shorter training time compared with BPNN. Based on data generated by a 3D Computational Fluid Dynamics (CFD) simulation, Iliyas et al. [11] developed a RBF neural network as soft sensor for high performance prediction of NO_X as well as O_2 of a 160 MW natural gas boiler. These researches illustrated that fast, agile and generalized learning of process model is really desired technology on advanced emission prediction.

General Regression Neural Network (GRNN) is a kind of RBF network that can perform fast learning and converge to the optimal regression surface quickly for large-scale samples [25]. GRNN has some attractive features for time sensitive applications: (1) it approximates any arbitrary function between inputs and outputs from sparse and noisy data; (2) it can be constructed quickly since only one iteration is needed to train an optimal network; (3) network design is simple since only one network parameter called smooth factor need to be set; (4) the estimation cannot converge to poor solutions corresponding to local minima of the error criterion. There have been numerous applications of GRNN in complex science and engineering process such as plasma microtrenching [12], batch chemical process [13], river sediment yield [14], water quality forecasting [15], public transportation [16], compressor performance prediction [17], electric power load forecasting ([18]). and prediction of sound absorption [19]. On prediction of boiler emissions, Zheng et al. [24] employed K-means clustering to adjust multiple smooth factors of GRNN for NO_X prediction. However, GRNN has not been investigated systematically in emission prediction, particularly on improving the on-line prediction quality and performance.

On the other hand, most existing works on emission prediction mainly focus on single emission such as NO_X, unburned carbon or SO_2 [5,7,8,10,11]. Prediction of multiple emissions has not gained widely attention, which can provide more information and more comprehensive understanding of boiler combustion to engineers. Using SVM for multiple characteristics prediction needs to construct several SVMs and set corresponding parameters respectively, which is not a simple solution for on-line application [9]. With the deeply understanding of environmental protection and the application of advanced equipment such as SCR, multiple emissions prediction will play more important role in the future's emission control.

Therefore, the motivation of this study is to investigate some improvements on GRNN for high quality on-line prediction of complex industrial process, particularly for multiple emissions prediction. As a result, an incremental learning method is employed for compact on-line GRNN modeling, then two efficient and fast methods, wiggle method and extrapolation node of GRNN, are proposed to enhance globally estimation ability of GRNN, and finally a simple and uniform framework regarding multiple emission prediction, on-line prediction, accurate prediction, fast prediction and self-learning ability is implemented.

2. Problem statement and contributions

In frequently changing industrial environment, fast response and adapting to external events is highly demanded for on-line application [26,27]. To build an advanced on-line multiple emissions prediction system based on GRNN, some issues should be addressed: first, how to perform fast on-line training and estimating for large-scale samples; second, how to overcome higher estimating error of GRNN in some input areas while keeping fast estimation; and third, how to efficiently estimate extreme value in on-line environment.

To this end, a globally enhanced GRNN neural network, GE-GRNN, is proposed to deal with these problems. Gaussian Adapted Resonance Theory (GART), an incremental learning method, is considered for training GRNN, which can control the memory cost by compressing the training samples into the hidden layer of GRNN. GART can perform learning on large-scale input samples and from the first sample [27,28]. Therefore, it is a good choice for on-line boiler emission modeling and prediction. Compared with existing works, this is the first one to introduce GRNN and GART learning to the task of on-line multiple emissions prediction. Since the estimation of GRNN is not accurate enough in some input areas (see Section 4 and Section 5.1), solutions are investigated considering global characteristics of the trained model along with on-line application, including selection of smooth factor and extrapolation.

The contributions of this work reflect in three aspects:

(1) modified wiggle-method is proposed to adjust smooth factors of GRNN in multi-dimension space to improve estimation accuracy considering both global and local characteristics of input samples;
(2) fast polynomial extrapolation structure is designed and added into the hidden layer of GRNN neural network to improve the quality of extreme value estimation;
(3) hybrid estimation mechanism is designed to integrate wiggle-method and extrapolation into an uniform estimation process, which makes the estimation more accurate and flexible.

The enhancements enable a uniform framework for on-line prediction, quickly prediction, accurately prediction and multiple emission prediction with fast self-learning ability, which provide a better solution to the problems in this study.

3. Related work

3.1. GRNN

GRNN is established on the theory of non-linear regression [25]:

$$\hat{Y}(\boldsymbol{X}) = E[y|\boldsymbol{X}] = \frac{\int_{-\infty}^{+\infty} y f(\boldsymbol{X}, y) dy}{\int_{-\infty}^{+\infty} f(\boldsymbol{X}, y) dy} \tag{1}$$

Where X is a particular measured vector of random variable x; y is scalar random variable; $\hat{\boldsymbol{Y}}(\boldsymbol{X})$ is conditional mean of y given X; and $\boldsymbol{f}(\boldsymbol{X}, \boldsymbol{y})$ is joint continuous probability density function. If $\boldsymbol{f}(\boldsymbol{X},$

Please cite this article as: J. Song et al., A globally enhanced general regression neural network for on-line multiple emissions prediction of utility boiler, Knowledge-Based Systems (2016), http://dx.doi.org/10.1016/j.knosys.2016.11.003