Monitoring and analysis of air emissions based on condition models derived from process history

M. Liukkonen and T. Hiltunen

Cogent Engineering (2016), 3: 1174182
COMPUTER SCIENCE | RESEARCH ARTICLE

Monitoring and analysis of air emissions based on condition models derived from process history

M. Liukkonen1* and T. Hiltunen2

Abstract: Evaluation of online information on operating conditions is necessary when reducing air emissions in energy plants. In this respect, automated monitoring and control are of primary concern, particularly in biomass combustion. As monitoring of emissions in power plants is ever more challenging because of low-grade fuels and fuel mixtures, new monitoring applications are needed to extract essential information from the large amount of measurement data. The management of emissions in energy boilers lacks economically efficient, fast, and competent computational systems that could support decision-making regarding the improvement of emission efficiency. In this paper, a novel emission monitoring platform based on the self-organizing map method is presented. The system is capable, not only of visualizing the prevailing status of the process and detecting problem situations (i.e. increased emission release rates), but also of analyzing these situations automatically and presenting factors potentially affecting them. The system is demonstrated using measurement data from an industrial circulating fluidized bed boiler fired by forest residue as the primary fuel and coal as the supporting fuel.

Subjects: Neural Networks; Power & Energy; Systems & Control Engineering

Keywords: circulating fluidized bed; energy boiler; self-organizing map; emission; monitoring

ABOUT THE AUTHOR

The University of Eastern Finland (UEF) enjoys a leading national position in the field of forestry research. The research projects of the Process Informatics research group in the Department of Environmental Science have focused on advanced measurement and modeling methods and advanced systems utilized by the process industry. M. Liukkonen, born in Jyväskylä, Finland, graduated from the University of Oulu, Finland, as MSc (eng., Process technology) in 2007 and from the University of Eastern Finland as PhD (Environmental technology) in 2010. His main research interests include analysis and modeling of measurement and other data, decision support and monitoring systems, advanced software applications, and advanced measurement systems. Research covers a variety of industrial environments including water and wastewater treatment, energy, pulp and paper, mining industries, and electronics production.

PUBLIC INTEREST STATEMENT

Air emissions are causing, not only deterioration of natural and cultural environment but also costs for the societies due to hospital visits, increased mortality and decrease in working years, and restoration of the environment. Therefore, air emissions are highly regulated, and it is obvious that the environmental legislation and regulations concerning energy plants will be tightened in the near future. Evaluation and analysis of online information about the process conditions is an essential task when reducing emissions in energy plants, and automated monitoring and control are of primary concern when preventing emissions, particularly in biomass combustion. Due to these challenges, we have developed a novel emission monitoring platform based on the self-organizing map method. The system is capable, not only of visualizing the prevailing status of the process and detecting problem situations (i.e. increased emission release rates) but also of analyzing these situations automatically and presenting factors potentially affecting them.
1. Introduction

Air emissions are currently highly regulated, and it is obvious that the environmental legislation and regulations concerning energy plants will be tightened in the near future. Despite these regulations, energy producers have to be able to respond quickly to altering load demands (Wang, Huang, Liu, Yan, & Cen, 2008) and tolerate variations in fuel quality (Beér, 2000), which emphasizes the importance of prompt and precise operational control of combustion. Furthermore, since fuel costs cover the majority of the operating costs of energy boilers (Basu, 1999), the increasing price of electricity creates additional challenges, and producers may be obliged to favor inexpensive, poor-quality fuels like waste, recycled fuels, and low-grade coal. The state of affairs is complicated, because varying fuel quality may have unpredictable effects on combustion and, on the other hand, process modifications may be impracticable or too expensive.

The flip side of the problematic solid fuels is their influence on the nature and on the amount of air emissions released. It is possible that due to a change in the fuel quality, the combustion conditions may temporarily become more favorable for the formation reactions of emissions, for which their formation is suddenly accelerated. Alternatively, low-grade fuels may include an enriched amount of fuel-bound emission precursors that are released and activated during combustion.

The technology of fluidized bed combustion (FBC) is especially suitable for converting solid fuels (coal, peat, biomass, etc.) into energy. In a circulating fluidized bed (CFB), the combustion occurs in a suspension (i.e. fluidized bed), which is typically a mixture of sand, fuel ash, and a possible sorbent material for sulfur capture. The bed material is fluidized by feeding in some air through the bottom of the combustion chamber (primary air). High fluidizing velocities are used in CFBs, so the bed particles are in a constant circulating movement within the system. Due to the large heat capacity of the bed, the combustion is generally stable, and supporting fuels such as oil or gas are usually needed only during the start-up. The intense turbulence of the CFB ensures good mixing and combustion. Combustion typically takes place at bed temperatures of ca. 850–900°C.

CFBs are extremely flexible with fuels, and they are able to process a wide variety of fuels without a noteworthy reduction in performance. This robustness enables co-combustion of a wide variety of solid fuels, which is a major advantage in the present energy market (Coda Zabetta, Hiltunen, Moulton, & Hotta, 2009). On the other hand, this can be considered a drawback as well, because changing fuel types and varying fuel quality typically inflict new challenges when it comes to controlled and stable combustion. For example, the use of alternative fuels, which are typically of lower quality, tends to worsen and complicate the NOx emission performance of combustion systems (Hill & Douglas Smoot, 2000). Sulfur dioxide (SO₂), nitrogen oxides (NOₓ), and nitrous oxide (N₂O) are the principal air pollutants released from CFB boilers fired by fossil fuels (Basu, 1999).

Assessment of online information on operating conditions is an essential task when reducing emissions in energy plants, and automated monitoring and control are of primary concern when preventing emissions, particularly in biomass combustion (Nussbaumer, 2003). The goal in traditional process monitoring is to detect abnormalities in process operation (MacGregor & Cinar, 2012), and the simplest ways of monitoring are to either observe the process performance by naked eye or to perceive the changes in it by using a preselected set of key measurements. Nevertheless, monitoring of industrial processes, including combustion and its emissions, is generally more demanding than this, which can be concluded from many facts. To begin with, emission monitoring with diagnosing abilities would be more useful than the general approach, because it is able to reveal the prevailing status of the process and present the potential reasons for possible problems. Moreover, automated computerized monitoring systems are usually needed by the industry for supporting decisions, because analyzing the prevailing process conditions based on an extensive data-set of measurements is arduous and difficult, even to a practiced specialist who has abundant former knowledge of the process. On the other hand, the utilization of automated monitoring systems makes it necessary to get reliable information on the process by systematizing and organizing measurements, which often leads to other dilemmas.
Principal component analysis and partial least squares are the customary methods of choice for monitoring systems with diagnosing abilities (MacGregor & Cinar, 2012; Qin, 2012). Nonetheless, when used in process monitoring, these two methods seem to be restricted by the lack of appropriate visualization properties for presenting the essential parts of extracted information in an understandable manner, such as the prevailing process status, and for analyzing the problem situations. As a result, the outputs of any systems based on these platforms are difficult and laborious to interpret. Monitoring based on a limited number of certain key measurements is sometimes feasible (Liukkonen, Juntunen, Laakso, & Hiltunen, 2013), but this kind of approach requires special knowledge of the process and is thus most applicable to cases in which the group of key variables is easily recognized and relatively small. Taken as a whole, there is room for new kinds of advanced emission monitoring and analysis systems.

The following topical issues related to modeling, management, and monitoring of NO\textsubscript{x} emission in CFB boilers can be identified:

• Monitoring of NO\textsubscript{x} in CFBs lacks economically efficient, fast, and competent computational systems that could support decision-making for improving the emission efficiency.

• A large part of the monitoring applications designed for CFB, especially those based on physical models, are case-specific and entail arduous model development and special knowledge on the process, and new cost-efficient alternatives to this would be useful.

• As monitoring of CFB emissions is ever more challenging because of low-grade fuels and fuel mixtures, new monitoring applications are needed to extract essential information from the large amount of measurement data.

• In these complicated industrial environments, it is essential to find out reasons for problems and to know the prevailing status of the process, and the methods currently available do not directly provide these kinds of functionalities.

Complex and dynamic industrial processes are difficult in terms of online monitoring. For example, condition-dependent behavior is characteristic for industrial processes (Juntunen, Liukkonen, Lehtola, & Hiltunen, 2013a, 2013b; Liukkonen, Havia, Leinonen, & Hiltunen, 2011; Liukkonen, Heikkinen, et al., 2011; Liukkonen, Hiltunen, Hälkkä, & Hiltunen, 2011), which usually requires flexibility and adaptivity from practical software applications. In practice, this means that the process adheres to different regularities under different conditions, which means that these conditions have to be recognized by the application, so that it can adapt to the prevailing situation. A typical example on this is load-dependent behavior in energy and water processes (Liukkonen, Heikkinen, et al., 2011; Liukkonen & Hiltunen, 2014; Liukkonen, Laakso, & Hiltunen, 2013). Moreover, it is shown that problems originating from fluctuation can be significant in a dynamically behaving process (Liukkonen, Hiltunen, et al., 2011). These may occur in transition circumstances such as load change situations, and inflict undesired side effects such as increased emission release rates. These kinds of challenges are frequently faced by energy producers, who have to deal with uncontrollable factors such as varying fuel quality, which may pose rapidly or slowly arising problems including fouling, corrosion, and air emissions. For these reasons, it would be useful to be able to observe the combustion performance and process status more efficiently.

Modeling and model-based predictive systems are potential approaches for improving the control of complex industrial processes (Keskitalo, la Cour Jansen, & Leiviskä, 2010; Liukkonen, Hälkkä, Hiltunen, & Hiltunen, 2012, 2013; Sanjay & Prithvi, 2014), and models based on process history (i.e. measurement data) are a potential approach to be used as the basis of advanced monitoring systems. For example, model-predictive monitoring of emissions has been of wider interest recently, and methods such as neural networks are widely used as the basis of predictive emission monitoring systems (Iliyas, Elshafei, Habib, & Adeniran, 2013; Lv, Liu, Yang, & Zeng, 2013; Shakil, Elshafei, Habib, & Maleki, 2009; Smrekar, Potočnik, & Senegačnik, 2013). A descriptive data-based system for monitoring and evaluating dynamic industrial processes has to be able to
process a large amount of multivariate data, extract the fundamental information from the data, and provide an understandable presentation of this information to be evaluated by experts. Additionally, these so-called intelligent monitoring platforms should be able to cope with multivariate and dynamic characteristics, which are typically inherent in industrial processes.

A promising platform for developing descriptive monitoring and diagnostics systems is the self-organizing map method (SOM) (Kohonen, 2001; Oja, Kaski, & Kohonen, 2002), its variants such as time adaptive SOM (Shah-Hosseini & Safabakhsh, 2000) and growing hierarchical SOM (Dittenbach, Merkl, & Rauber, 2000), the combination of SOM and k-means clustering (Chang & Chen, 2015; Liukkonen, Heikkinen, et al., 2011; Liukkonen, Laakso, et al., 2013), and the alternatives of SOM such as generative topographic maps (Bishop, Svensén, & Williams, 1998) and elastic maps (Gorban & Zinovyev, 2005).

The basic SOM is an artificial neural network based on unsupervised learning. The training of SOM results in a topological arrangement (lattice), of output neurons, each of which has a special property vector describing the input vectors associated with it. During training, these so-called reference vectors become weighted averages of the original data samples assimilated to it. In the lattice, the input vectors sharing common features end up into the same or neighboring neurons. As a result, the grid of neurons (i.e. the map) reflects variations in the statistics of the data-sets and selects common features, which approximate to the distribution of the data points. Recently, the SOM has been widely used in a wide variety of environmental monitoring systems (Chea, Grenouillet, & Lek, 2016; Strebel et al., 2013; Yu et al., 2014; Zhang, Liang, Wang, & Xu, 2015).

Nonetheless, the methods mentioned above do not provide a direct way of monitoring and diagnosing industrial processes as such. In the earlier study (Liukkonen & Hiltunen, 2014) we presented an approach for advanced monitoring of emissions in CFB boilers. In this approach, a model based on SOM and fuzzy c-means clusters is updated regularly to respond to the prevailing condition, the current condition of the process is monitored and evaluated, and an alert is presented whenever the preset limit for each emission component is exceeded. Although found suitable for monitoring emissions in a way that is not dependent on the load level, the method lacks the basic functionality required in online monitoring applications, and the structure based on clustering the map after every training stage may be unnecessarily complicated and heavy when it comes to flexible monitoring and simple visualization platforms. Thirdly, the approach is not fully capable of presenting an efficient diagnosis of the problem at hand, because the analysis is performed using the cluster centers, and not the prototype vectors of the model itself.

Keeping these limitations in mind, the objectives of this study were

- to determine if SOM-based models derived from process history could be applicable to automated, online emission monitoring systems,
- to design a flexible, adaptive, and visual decision-support system based on data-driven models for the improvement of emission efficiency in energy boilers,
- and to provide the system with automatable diagnostic abilities in problem situations.

The novelty of the SOM-based platform presented in the paper becomes from the fact that it can be utilized for decision-support in automated emission monitoring applications. It is also more practical in visualizing the dynamics and evolution of the process, and the structure of the platform is more simple, because the map units are used as conditional prototypes when determining the present condition of the process. Furthermore, the system is capable of presenting a diagnosis whenever a problem, i.e. an increase in the emission release rate, is detected. The environment in which the system is simulated involves measurement data from an industrial CFB boiler fired by forest residue as the primary fuel and coal as the supporting fuel.
2. Methods and data

2.1. The basis of the adaptive emission monitoring system

Self-organizing maps (Kohonen, 2001) are a well-known, unsupervised form of artificial neural networks that is broadly used in descriptive data analysis (Oja et al., 2002). The following phases are involved in the training procedure of a single SOM:

1. Initialization of the reference vectors of neurons
2. Searching for a best matching unit (BMU) for the first input vector based on Euclidean distance metrics
3. Recalculating the reference vector of the BMU based on the values of the input vector
4. Recalculating the reference vectors of the neurons located next to BMU
5. Iteration of steps 2–4 for every input vector in sequence
6. Fine tuning of the map by iterating steps 2–5 using a smaller learning rate
7. Searching for the eventual BMUs for the input vectors

In other words, \( n \)-dimensional data are assigned to structural units (i.e. neurons), which are usually hexagonal and which can be arranged in a multidimensional lattice (i.e. map). Generally, a two-dimensional map is used, because it can be visualized and interpreted more easily. In the supervised training procedure, every neuron on the map adapts to two characteristic features: (1) the position on the lattice, and (2) an \( n \)-dimensional reference vector. The first one of these features describes the degree of similarity of the neuron to other neurons (the closer the neurons are to each other, the more similar they are), whereas the other one presents a description of the archetypal features of the original data samples assigned to this neuron. Therefore, the reference vectors can be considered, on one hand, as generalized descriptions of the original data vectors assigned to each unit, and on the other hand, as estimators of the similarity of neurons.

The drawback of the traditionally used visualization methods of SOM is that they do not offer a direct way of monitoring the process dynamics and evolution. Another visualization approach for use with SOM is the trajectory analysis, which can be used for illustrating the short- and long-term behavior of a dynamical system and the mechanisms by which the system is changing with time. A SOM trajectory demonstrates the order of BMUs in the SOM in response to a set of input vectors taken over a given period of time (Kohonen, Oja, Simula, Visa, & Kangas, 1996). Generally speaking, trajectory points are first acquired by determining the BMU of each original data sample, and these points can then be illustrated successively on the map. In practice, trajectories can thus be formed using consecutive images of operating points that can be displayed on the map after training (Kohonen, 2001).

2.2. Dynamical monitoring approach

The SOM-based system for adaptive emission monitoring is organized as follows:

(a) Train the SOM using recorded measurement data archives to create the background model to be used by the system.
(b) Input the limits of variables to be monitored (only when implemented for the first time or when these limits are changed).
(c) Feed input data consisting of the most recent measurements (consistent feed).
(d) Obtain output data including the process evolution (i.e. trajectories) and the most recent conditions (i.e. neurons) to be shown on the map, as well as alerts and information on possible emission problems, which are also reported and analyzed.
The main stages of the adaptive monitoring approach are illustrated in Figure 1, and a diagram describing the operation of the system in a decision-making task is presented in Figure 2. Right after feeding in the measurement data obtained at a single moment of time, including the emission components to be monitored (NO\textsubscript{x} and SO\textsubscript{2} emissions), to the system, the BMU of this data row is first determined by terms of Euclidean distance. Using this information and a number of most recent BMUs, the present and former conditions can be shown as a trajectory on the SOM. Generally speaking, the application determines the prevailing condition of the process by finding a model (i.e. the reference vector of a neuron) which is the most similar to the most recent measurement.
After this, the recently measured emission (NOX and SO2) release rates are compared to the preset limits to analyze if they are exceeded. An exceeding of the red limit for an effluent component leads to a notification and an alarm sound, and in the meantime a red circle is shown on the map. In case there is an exceeding of the yellow warning level, the color tone of the circle is extracted from a color map comprising tones from yellow to red, depending on the degree of the exceeding. As a consequence, the colored circle appearing on the map illustrates, not only the problem itself but also the severity of the problem.

This step is followed by a diagnosis which eventually proposes the potential reasons for the problem situation. First, the reference vector of the present BMU is analyzed by the system to search for those variable components that have the largest absolute values. As a result, those variable components which distinguish this neuron the most from other neurons can be illustrated on the screen, and this way the possible reasons for the problem become achievable. Furthermore, the latest values of these variables are shown by the graph, which enables the comparison between the most recent measurements and the reference vector. This makes it possible to evaluate how similar the prevailing conditions are to the prototype conditions to which they have been assimilated.

### 2.3. Diagnosis of high emission rates using condition analysis

As reference vectors provide generalized descriptions of the original data vectors assigned to each unit, each neuron represents a model derived from a number of data vectors that are closely related to each other in terms of Euclidean distance. These models can be regarded as condition models illustrating the prevailing state of the process.

For example, let $r_m$ be an $n$-dimensional reference vector describing the condition of a SOM unit:

$$r_m = (r_{m1}, r_{m2}, \ldots, r_{mn}), (m = 1, \ldots, M)$$
where $n$ is the number of model variables and $M$ refers to the total number of map neurons. Due to the nature of SOM, vectors $r_1$ and $r_2$ are quite similar as well as vectors $r_{107}$ and $r_{108}$, whereas vectors $r_1$ and $r_{107}$ are significantly different, as can be seen in Figure 3. For the same reason, reference vectors may also present the potential reasons for any occurred problems: if the problem is detected in a situation that most likely belongs to the neuron having the reference vector $r_1$, the reasons have to be different than those of a problem detected in a situation that belongs to the neuron of reference vector $r_{107}$. In other words, the factors that separate these conditions from the other conditions are important.

### 2.4. Case: Fluidized bed process and data

FBC technology is intended primarily for producing energy from solid fuels such as coal, peat, biomass, or waste-derived fuels. CFBs, an example of which can be seen in Figure 4, are a widely used form of FBC, which are able to burn a large variety of fuels without a major reduction in performance (Koornneef, Junginger, & Faaij, 2007). This flexibility allows co-combustion of solid fuels, including fossil fuels such as coal, biomass-derived fuels such as wood and agricultural residues, peat, and even waste-derived fuels like demolition wood.

The demonstration environment involves a 150 MWth industrial CFB boiler fired by forest residue (mainly bark) as the primary fuel and coal as the supporting fuel. The total number of samples is 8,640 from a period of one month of operation (5-min sample interval). The total number of input variables is 57 (see Table 1).

### 3. Case demonstrations

#### 3.1. Monitoring of NO$_x$ and SO$_x$ in 150 MWth CFB

The emission monitor based on SOM-derived condition models is coded on the Matlab (Mathworks, Natick, USA) platform. To demonstrate the use of the system, we present two examples of problem situations on increasing emission rates. An example in which an increased NO$_x$ concentration is detected can be seen in Figure 5. In the component planes (i.e. the colored maps), the colors of individual hexagonal neurons varying from red to blue indicate the value of the variable concerned: a red color tone on the SOM indicates that the value of the variable, in general, is on a high level in that neuron, whereas a blue color tone denotes a low value of that variable.

Thus, it can be seen in the large map that the data examples associated with a high level of main steam flow (i.e. the boiler load level) have been concentrated on the lower left corner of the map. The trajectory shown on the leftmost map illustrates that the process evolves from a state of medium power to that of high power, whereby also the NO$_x$ emission rate is increased. Meanwhile, warnings are presented in the right-hand list. Finally, an alarm is given as the NO$_x$ emission continues to increase.
A condition analysis based on the reference vectors is then performed and presented as a bar graph shown on the right in Figure 5. As can be seen in the graph, this situation is characterized by a high boiler load. Moreover, it is shown that the bark to coal ratio is high, meaning that the share of bark is high compared to the share of coal. Perhaps there is a set of biomass fuel in use that has a high content of nitrogen. In addition, the bed pressure seems to be low in this problem situation.

Another problem situation in which an increased SO₂ concentration is detected can be seen in Figure 6. As can be seen, the power of the boiler is on the low level in this case. It is also notable that in this situation, the boiler is fueled mainly by the supporting fuel (i.e. coal), which is not the preferable fuel combination based on which the boiler was designed. In summary, low boiler load and perhaps relatively unstable combustion conditions, added to the use of a sulfur-rich fuel, has led to a release of a relatively high concentration of sulfur dioxide.

### Table 1. The variables of sample data

| Variable                          |Measure-ments¹ |
|-----------------------------------|----------------|
| (1) Fuel and sand feed            |                |
| Fuel screw speed (bark)           | 2              |
| Fuel screw speed (coal)           | 2              |
| Bark/coal ratio                   | 1              |
| Total bark flow                   | 2              |
| Fuel feed (bark)                  | 1              |
| Fuel feed (coal)                  | 1              |
| Flow of start-up oil              | 1              |
| Bark moisture                     | 1              |
| (2) Air and gas feed              |                |
| Primary air flow                  | 1              |
| Secondary air flows               | 4              |
| Secondary air flow (total)        | 1              |
| Flow of circulating gas           | 1              |
| Primary air temperature after fan | 1              |
| Secondary air temperature after fan| 1             |
| (3) Bed and furnace conditions    |                |
| Bed pressure                      | 3              |
| Furnace pressure                  | 5              |
| Bed temperature                   | 5              |
| Bed temperature (avg)             | 1              |
| Furnace temperature               | 3              |
| (4) Steam system                  |                |
| Main steam flow                   | 1              |
| (5) Flue gas                      |                |
| O₂, flue gas                      | 3              |
| CO, flue gas                      | 1              |
| SO₂, flue gas                     | 1              |
| NO₂, flue gas                     | 1              |
| H₂O, flue gas                     | 1              |
| Flue gas temperature              | 5              |
| Cyclone temperature               | 2              |

¹Number of measurements for each variable.
3.2. Analysis of problem situations

In addition to online performance monitoring, another useful feature of the approach is the offline (post-process) analysis of emissions. The summary of problem situations in the CFB case can be seen in Figure 7. It can be seen that the problems related to a high SO2 concentration are concentrated in two quite similar process states (i.e. neurons), illustrated on the bottom right corner of the map. The combination of a low boiler load and a high coal to bark ratio are characteristic features for these neurons.

The situation is essentially different when it comes to problems related to a high NOx concentration, however. First of all, the NOx-related problems can be linked with all degrees of boiler load. Under low load conditions, the temperatures of feed air, a high bed pressure, and a high coal to bark ratio are the potential reasons for increased NOx release rate, whereas high temperatures and a low bed pressure are characteristics for the problem situations detected under medium load conditions. The latter cases (medium power) seem to be situations in which the load is presently being reduced, and the process has not yet gained its stability. In other words, these factors indicate that there has been some instability in the process. When it comes to high boiler load, it is obvious that some of the problems originate from the fact that the process is working at full power and is therefore...
approaching the unacceptable emission limits. On the other hand, the low bed pressure may be a sign of a disturbance in the combustion, which has led to an exceptionally increased NOx concentration in the flue gas.

4. Discussion

Environmental awareness has been increased exponentially in recent decades, and it is clear that power plants will have to seek new ways of managing emissions more efficiently in the future. Measurement databases are valuable process history records, because they contain information on the past operation of the process and on the mutual interactions between different factors affecting the process performance. When it comes to emission management, these data can be exploited in several tasks, such as estimating future emission concentrations or analyzing the emissions already released. This kind of refined information also provides support for decisions aimed at reducing the emissions in the future. Methodologies based on computational intelligence, including intelligent monitoring systems and soft sensors, offer a promising alternative route to process improvement, by which a more efficient and environmentally friendly process can be achieved.

As industrial processes are getting more complex and the amount of measurements is increasing continuously, it is obvious that process monitoring based on single measurements is not the most efficient way of controlling the processes anymore. Monitoring systems having diagnosing abilities and based on multivariable measurement data can provide mutual information that provides deeper knowledge and is thus more useful, but this kind of information is accessible only by models which are able to extract the essential features from the data. On the other hand, the outputs of the multivariable process monitoring methods used traditionally are not very easy to interpret, as they lack holistic visualization abilities for presenting the extracted information such as the prevailing condition of the process and its direction. What is even more important, their capability of analyzing problem situations and their potential reasons is generally restricted, because they are designed principally to detect normal conditions.

4.1. Interpretation of results

In the present paper, an adaptive monitoring platform based on the SOM method is described, and the system is demonstrated using measurement data from an industrial CFB boiler fired by forest residue as the primary fuel and coal as the supporting fuel. The system is capable, not only of visualizing the process behavior and detecting the problem situations but also of providing analysis of
these cases automatically. Based on the results, the novel method for monitoring emissions over-
comes the existing ones in certain aspects that should be important in future monitoring applica-
tions. The pros and cons of the most general multivariate methods with diagnosing abilities for
emission monitoring can be seen in Table 2.

The chemical and physical phenomena occurring in combustion are typically dynamical, which
has to be taken into account when developing monitoring systems. The present results show that an
adaptive approach able to learn the changing behavior is necessary to adapt to the dynamics of the
CFB combustion. It can also be concluded that this kind of dynamical evolution can be compre-
hended more easily by monitoring the change between different conditions, i.e. the transition from
certain identified status to another one. This can be described by a model constructed using a group
of key measurements and by visualizing this model in an understandable manner.

This, combined with the presented results, leads us to the most important conclusion, which is
that the SOM-based adaptive monitoring approach demonstrated here offers a descriptive, visual
way of modern emission monitoring and analyzing both immediate changes and long-term evolu-
tion in CFB combustion. The model behind the visual presentation is able to learn with time, which
makes it capable of detecting new, previously unobserved phenomena. The system also facilitates
the extraction of information from a large amount of multidimensional measurement data and can
be used for comprehensive, model-based condition monitoring.

4.2. Challenges of condition monitoring

Because manual processing of measurement data tends to be laborious and time consuming, ad-
vanced data mining methods can provide resource savings in the monitoring, diagnosis, and even
control of industrial processes. The so-called “big data” problem, i.e. the continuously increasing
amount of measurement data, and increasingly complicated processes require appropriate soft-
ware tools to extract the essential information. For the same reasons, intelligent software systems
designed for advanced process diagnostics can prove themselves useful. The results of this study
provide a framework that can be easily implemented in a software application. This kind of software
provides a competent way of analyzing a large amount of process data, and the results show that it
could provide a useful modeling tool for industrial applications.

However, it is doubtful that systems such as presented here can be completely autonomous in the
near future, and it is presumable that human intervention is needed in certain steps of evaluation.
For example, in this case the model working in the background of the system also involves the

| Attribute                                      | PCA | PLS | SOM |
|------------------------------------------------|-----|-----|-----|
| Ability to detect abnormalities                | ++  | ++  | ++  |
| Ability to use multivariate data               | ++  | ++  | ++  |
| Interpretation of results                      | +   | +   | +   |
| Nonlinearity                                   | *   | *   | ++  |
| Visualization of transitions                   | +   | +   | +   |
| Ability to provide status information          | +   | +   | +   |
| Ability to diagnose problems                   | +   | +   | ++  |
| Ability to learn                               | ++  | ++  | +   |
| Requirements for data quality                  | ++  | ++  | +   |
| Ease of use (without a tailored solution)      | +   | +   | +   |
| Commercial availability                        | ++  | +   | +   |

Notes: PCA = principal component analysis, PLS = partial least squares, SOM = self-organizing map.
*Nonlinear version exists also.
functionality of analyzing those situations which are classified as problematic. Nonetheless, it is crucial to be aware of the fact that the system makes, in principle, just a preliminary suggestion on the possible reasons for the problem events. These potentially affecting factors can be viewed on the screen and have to be reviewed by an expert who has adequate experience on the process performance and the required empirical knowledge to judge which ones of the presented variables are the most logical origins of this particular problem. This evaluation is critical, because the variables brought up by the system are simply a mathematical representation on the factors by which the prevailing process state differs most from the others.

The requirement for proper training data is another important issue when implementing a system such as presented here. It is important to bear in mind that the ability to adapt to prevailing conditions (i.e. the most recent data) is a prerequisite when applying the system in dynamical processes. Therefore, the model has to be kept up-to-date by training the SOM regularly using a representative sample data-set. On the other hand, in some cases it might be useful to preserve also some older information, and not just the most recent. For example, some phenomena affecting emission rates may happen regularly but rarely, so the system should be able to remember some older occurrences as well. Additionally, if the reason for a detected problem or a combination of reasons is not represented in the training data-set (the origins of the problem are new to the system), the system will not be able to evaluate the situation.

4.3. Recommendations and future aspects
Based on the results, it is possible to recommend the most beneficial purposes of use for the approach presented in this study. When it comes to process design and development towards cleaner processes, it is obvious that physical emission models will be the standard parts in the emission modeling toolbox also in the future. Therefore, the most practical use of advanced data-driven methods would be in supporting the physical emission models. A promising purpose of use for data-driven emission models is in advanced monitoring systems. The main advantages of these systems are that they facilitate the extraction of information from multidimensional measurement data, assist in interpreting the prevailing process status, and may describe its evolution in a comprehensive manner. The ability to adapt to prevailing conditions (i.e. the most recent measurement data) is a prerequisite when applying these methods in energy systems. The most practical way of using these methods would be in advanced software solutions for decision-support which are designed to diagnose and monitor industrial processes.

The most promising target of use of the presented approach is in decision support systems aimed at reducing air emissions and their management cost in energy boilers. In the future, an advanced emission monitoring system should be able to propose actions for improving the emission performance, so that the most problematic situations could be prevented by controlling the process beforehand. An intelligent monitoring system could be able to analyze the status of the process and assist the plant personnel in interpreting the prevailing risk level using the available measurements. The system could then suggest a code of practice, so that the process operator would be able to act correctly and quickly. It would also be important to know the future condition of the process and the future emission concentration, to anticipate the future status and to justify operative actions to improve and optimize the process. This could be achieved by a combination of intelligent data-based monitoring and soft sensors.

5. Conclusions
We have presented a new approach for advanced, model-based emission monitoring in CFBs, which is realized by combining existing methodology and multivariate data consisting of process measurements in a novel way. A real-world CFB process is presented as a case study. It can be concluded that data-driven emission models can provide functional platforms for advanced monitoring systems.

Based on the results the main advantages of the model-based monitoring system are as follows:
• Good descriptive capabilities including visualization of short- and long-term changes.
• Ability to analyze: the system is able to diagnose problem situations, recognize prevailing process status and provide decision-support for corrective actions.
• Ability to adapt: the system is able to adapt to prevailing conditions.
• Ability to learn: the system is able to learn from process history.

As the industrial monitoring platforms are moving towards more intelligent solutions which provide, not only the traditional detection of abnormal conditions but also many predictive and analytical capabilities, the systems such as presented here offer potential platforms for this purpose. As the regulations regarding the industrial emissions will be tightened in the near future, the results serve the development of advanced decision-support systems intended to reduce air emissions.

Acknowledgements
The Finnish Funding Agency for Technology and Innovation (TEKES) and the project Intelligent Software and Service Concept of the Industrial Internet (InDiGO!) are acknowledged.

Funding
The Finnish Cultural Foundation and the Centenary Fund of Pohjola Companies are highly appreciated for funding the research.

Author details
M. Liukkonen1
E-mail: mika.liukkonen@uef.fi
T. Hiltunen2
E-mail: teri.hiltunen@fwfin.fwc.com
1 Department of Environmental Science, University of Eastern Finland, P.O. Box 1627, Kuopio FIN-70211, Finland.
2 Foster Wheeler Power Group, P.O. Box 201, Varkaus FIN-78201, Finland.

Citation information
Cite this article as: Monitoring and analysis of air emissions based on condition models derived from process history, M. Liukkonen & T. Hiltunen, Cogent Engineering (2016), 3: 1174182.

Cover image
Source: Miko Liukkonen (2016).

References
Basu, P. (1999). Combustion of coal in circulating fluidized-bed boilers: A review. Chemical Engineering Science, 54, 5547–5557. http://dx.doi.org/10.1016/S0009-2509(99)00285-7
Beér, J. M. (2000). Combustion technology developments in power generation in response to environmental challenges. Progress in Energy and Combustion Science, 26, 301–327. http://dx.doi.org/10.1016/S0360-1285(00)00007-1
Bishop, C. M., Svensén, M., & Williams, C. K. I. (1998). GTM: The generative topographic mapping. Neural Computation, 10, 215–234. http://dx.doi.org/10.1162/089976698300017953
Chang, C. C., & Chen, S. H. (2015). A comparative analysis on artificial neural network-based two-stage clustering. Cogent Engineering, 2, 995785. http://dx.doi.org/10.1080/23311916.2014.995785
Chea, R., Grenouillet, G., & Lek, S. (2010). Evidence of water quality degradation in lower Mekong basin revealed by self-organizing maps. PloS One, 11, e0145527. http://dx.doi.org/10.1371/journal.pone.0145527
Coda Zabetta, E., Hiltunen, M., Moulton, B., & Hotta, A. (2009). Biomass and waste co combustion – European experience. In Proceedings of Electric Power 2009 Conference. Rosemont, IL.
Dittenbach, M., Merki, D., & Rauber, A. (2000). The growing hierarchical self-organizing map. In Proceedings of International Joint Conference on Neural Networks (IJCNN 2000) (Vol. VI, pp. 15–19). Los Alamitos, CA: IEEE Computer Society.
Gorban, A. N., & Zinovyev, A. (2005). Elastic principal graphs and manifolds and their practical applications. Computing, 75, 359–379. http://dx.doi.org/10.1007/s00607-005-0122-6
Hill, S. C., & Douglas Smoot, L. (2000). Modeling of nitrogen oxides formation and destruction in combustion systems. Progress in Energy and Combustion Science, 26, 417–458. http://dx.doi.org/10.1016/S0360-1285(00)000011-3
Iliyas, S. A., Elshafei, M., Habib, M. A., & Adeniran, A. A. (2013). RBF neural network inferential sensor for process emission monitoring. Control Engineering Practice, 21, 962–970. http://dx.doi.org/10.1016/j.conengprac.2013.01.007
Juntunen, P., Liukkonen, M., Lehtola, M., & Hiltunen, Y. (2013a). Cluster analysis by self-organizing maps: An application to the modelling of water quality in a treatment process. Applied Soft Computing, 13, 3191–3196. http://dx.doi.org/10.1016/j.asoc.2013.01.027
Juntunen, P., Liukkonen, M., Lehtola, M. J., & Hiltunen, Y. (2013b). Dynamic soft sensors for detecting factors affecting turbidity in drinking water. Journal of Hydroinformatics, 15, 416–426. http://dx.doi.org/10.2166/hydro.2012.052
Keskitalo, J., J. Cour Jansen, J., & Leiviskä, K. (2010). Calibration and validation of a modified ASM1 using long-term simulation of a full-scale pulp mill wastewater treatment plant. Environmental Technology, 31, 555–566. http://dx.doi.org/10.1080/09593330903563802
Kohen, T. (2001). Self-organizing maps (3rd ed.). New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-56927-2
Kohen, T., Oja, E., Simula, O., Visa, A., & Kangas, J. (1996). Engineering applications of the self-organizing map. Proceedings of the IEEE, 84, 1358–1384. http://dx.doi.org/10.1109/5.537105
Koormneef, J., Junginger, M., & Faaii, A. (2007). Development of fluidized bed combustion – An overview of trends, performance and cost. Progress in Energy and Combustion Science, 33, 19–55. http://dx.doi.org/10.1016/j.pecs.2006.07.001
Liukkonen, M., Havia, E., Leinonen, H., & Hiltunen, Y. (2011). Quality-oriented optimization of wave soldering process by using self-organizing maps. Applied Soft Computing, 11, 214–220. http://dx.doi.org/10.1016/j.asoc.2009.11.011
Liukkonen, M., Heikkinen, M., Hiltunen, T., Hällikä, E., Kuivalainen, R., & Hiltunen, Y. (2011). Artificial neural networks for analysis of process states in fluidized bed combustion. Energy, 36, 339–347. http://dx.doi.org/10.1016/j.energy.2010.02.007
Liukkonen, M., Hiltunen, T., Hällikä, E., & Hiltunen, Y. (2011). Modeling of the fluidized bed combustion process and NOx emissions using self-organizing maps: An application
to the diagnosis of process states. Environmental Modelling & Software, 26, 605–614.
Liukkonen, M., Hälikkä, E., Hiltunen, T., & Hiltunen, Y. (2012). Dynamic soft sensors for NOx emissions in a circulating fluidized bed boiler. Applied Energy, 97, 483–490. http://dx.doi.org/10.1016/j.apenergy.2012.01.074
Liukkonen, M., Hälikkä, E., Hiltunen, T., & Hiltunen, Y. (2013). Adaptive soft sensor for fluidized bed quality: Applications to combustion of biomass. Fuel Processing Technology, 105, 46–51. http://dx.doi.org/10.1016/j.fuproc.2011.06.015
Liukkonen, M., Juntunen, P., Laakso, I., & Hiltunen, Y. (2013). Advanced monitoring of industrial wastewater treatment: Multivariable approach using the self-organizing map. Environmental Modelling & Software, 48, 193–201.
Liukkonen, M., & Hiltunen, T. (2014). Adaptive monitoring of emissions in energy boilers using self-organizing maps: An application to a biomass-fired CFB (circulating fluidized bed). Energy, 73, 443–452. http://dx.doi.org/10.1016/j.energy.2014.06.034
Lv, Y., Liu, J., Yang, T., & Zeng, D. (2013). A novel least squares support vector machine ensemble model for NOx emission prediction of a coal-fired boiler. Energy, 55, 319–329. http://dx.doi.org/10.1016/j.energy.2013.02.062
MacGregor, J. A., & Chin, A. (2012). Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven methods. Computers and Chemical Engineering, 47, 111–120. http://dx.doi.org/10.1016/j.compchemeng.2012.06.017
Nussbaumer, T. (2003). Combustion and co-combustion of biomass: Fundamentals, technologies, and primary measures for emission reduction. Energy & Fuels, 17, 1510–1521.
Oja, M., Kaski, S., & Kohonen, T. (2002). Bibliography of self-organizing map (SOM) papers: 1998-2001 Addendum. Neural Computing Surveys, 3, 1–156.
Qin, S. J. (2012). Survey on data-driven industrial process monitoring and diagnosis. Annual Reviews in Control, 36, 220–234. http://dx.doi.org/10.1016/j.arcontrol.2012.09.004
Sanjay, C., & Pritivi, C. (2014). Hybrid intelligence systems and artificial neural network (ANN) approach for modeling of surface roughness in drilling. Cogent Engineering, 1, 943935. http://dx.doi.org/10.1080/23311916.2014.943935
Shah-Hosseini, H., & Safabakhsh, R. (2000). TASOM: The time adaptive self-organizing map. In Proceedings of International Conference on Information Technology: Coding and Computing (pp. 422–427). Los Alamitos, CA: IEEE Computer Society.
Shakil, M., Eshkafelti, M., Habib, M. A., & Maleki, F. A. (2009). Soft sensor for NO and O, using dynamic neural networks. Computers and Electrical Engineering, 35, 578–586. http://dx.doi.org/10.1016/j.compeleceng.2008.08.007
Smrekar, J., Potočnik, P., & Senegačnik, A. (2013). Multi-step-ahead prediction of NOx emissions for a coal-based boiler. Applied Energy, 106, 89–99. http://dx.doi.org/10.1016/j.apenergy.2012.10.056
Strebel, K., Espinosa, G., Giralt, F., Kindler, A., Rallo, R., Richter, M., & Schlink, U. (2013). Modeling airborne benzene in space and time with self-organizing maps and Bayesian techniques. Environmental Modelling & Software, 41, 151–162.
Wang, F., Huang, Q., Liu, D., Yan, J., & Cen, K. (2008). Improvement of load-following capacity based on the flame radiation intensity signal in a power plant. Energy Fuels, 22, 1731–1738. http://dx.doi.org/10.1021/ef700620n
Yu, H., Song, Y., Liu, R., Pan, H., Xiang, L., & Qian, F. (2014). Identifying changes in dissolved organic matter content and characteristics by fluorescence spectroscopy coupled with self-organizing map and classification and regression tree analysis during wastewater treatment. Chemosphere, 113, 79–86. http://dx.doi.org/10.1016/j.chemosphere.2014.04.020
Zhang, Y., Liang, X., Wang, Z., & Xu, L. (2015). A novel approach combining self-organizing map and parallel factor analysis for monitoring water quality of watersheds under non-point source pollution. Scientific Reports, 5, 16079. http://dx.doi.org/10.1038/srep16079

© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:
Share — copy and redistribute the material in any medium or format.
Adapt — remix, transform, and build upon the material for any purpose, even commercially.
The licensor cannot revoke these freedoms as long as you follow the license terms.
Under the following terms:
Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.
You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
No additional restrictions
You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Cogent Engineering (ISSN: 2331-1916) is published by Cogent OA, part of Taylor & Francis Group.
Publishing with Cogent OA ensures:
• Immediate, universal access to your article on publication
• High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
• Download and citation statistics for your article
• Rapid online publication
• Input from, and dialog with, expert editors and editorial boards
• Retention of full copyright of your article
• Guaranteed legacy preservation of your article
• Discounts and waivers for authors in developing regions
Submit your manuscript to a Cogent OA journal at www.CogentOA.com