Nonlinear Prediction of the Hot Metal Silicon Content in the Blast Furnace

Henrik SAXÉN and Frank PETTERSSON

Heat Engineering Laboratory, Åbo Akademi University, Biskopsg. 8, FIN-20500 Åbo, Finland.
E-mail: henrik.saxen@abo.fi, frank.pettersson@abo.fi

(Received on June 14, 2007; accepted on October 3, 2007)

The processes in metallurgical industry are often extremely complex and measurements from their interior are scarce due to hostile conditions. Today's constraints on high productivity and minor impact on the environment still require that the processes be strictly controlled. Mathematical models can play a central role in achieving these goals. In cases where it is not possible, or economically feasible, to develop a mechanistic model of a process, an alternative is to use a data-driven approach, where a black-box model is built on historical process data. Feedforward neural networks have become popular nonlinear modeling tools for this purpose, but the selection of relevant inputs and appropriate network structure are still challenging tasks. The work presented in this paper tackles these problems in the development of a model of the blast furnace hot metal silicon content. A pruning algorithm is applied to find relevant inputs and their time lags, as well as an appropriate network connectivity, for solving the given time-series problem. In applying the model, an on-line learning of the upper-layer weights is proposed to adapt the model to changes in the input–output relations. The analysis shows results in good agreement with findings by other investigators and practical metallurgical knowledge.

KEY WORDS: black-box modeling; blast furnace; silicon prediction; neural networks; selection of inputs.

1. Introduction

During the last decades, neural networks have, inspired by their universal approximation capabilities, become popular tools in nonlinear black-box modeling. However, in real-world applications, such as prediction tasks encountered in the metallurgical industry, several practical problems arise in their use. A common problem is that there is an abundance of variables that potentially influence the output (dependent) variable, and a choice among these has to be made to avoid model over-parameterization. The presence of measurement errors also complicates the choice of network complexity, and the number of network parameters has usually to be restricted to avoid fitting noise. To address these problems, constructive and destructive algorithms, with growing or shrinking networks, have been proposed. A potential weakness of the constructive approaches is their sensitivity to noise, while the destructive approaches often include retraining steps that require prohibitive computational efforts. This paper illustrates how a novel pruning algorithm recently developed by the authors of the present paper can be used to tackle the complex problem of predicting the silicon content in hot metal produced in a blast furnace. This important problem has over the years been tackled by several investigators and has also acted as a benchmark problem for time-series modeling techniques. The proposed algorithm is here applied to detect both the relevant inputs and their time lags and furthermore to suggest a parsimonious model, i.e., proper network size and connectivity between the inputs and the hidden nodes. The algorithm is briefly outlined in the next section, including an extension of it to on-line learning. Section 3 first presents the prediction problem, and next demonstrates that the algorithm yields good solutions. A statistical analysis of the results of several runs of the algorithm is finally used to find a final model. The last section gives concluding remarks and proposes future directions of work.

2. The Pruning Algorithm

The main motivation for using pruning algorithms in modeling with neural networks is that the flexibility of the network models in combination with their large number of parameters (weights) make it necessary to control the complexity of the models to avoid over-fitting. Typically, connections or hidden nodes are pruned by gradually removing unnecessary ones until the results (e.g., the function approximation provided by the network) suffer. The pruning algorithm used in the present work is based on feedforward neural networks of multi-layer perceptron type (often also referred to as backpropagation networks) with a single layer of hidden nonlinear units and a single linear output node. Given a sufficient number of hidden nodes, such a network has been shown to be able to approximate any continuous, twice differentiable function to any accuracy. Fundamentally, the algorithm is based on the practical experience that for such a network with an arbitrary choice of weights in its lower layer of connections, W, (cf. Fig. 1) there is usually a weight vector, v, to the output node that...
will lead to a relatively good solution, \( \hat{y} \), of the approximation problem. For an example, the reader is referred to Saxén and Pettersson. \(^7\)

The vector \( \mathbf{v} \) can be determined by a simple matrix inversion. With this as the starting point, the pruning algorithm proposed can be compactly written as:

1) Select a set of \( N \) potential inputs, \( \mathbf{x} \), and the output, \( y \), to be estimated for the observations of the training set.
2) Choose a sufficient number of hidden nodes, \( m \), and generate a random weight matrix, \( \mathbf{W}^{(0)} \), for the lower part of the network. Set the iteration index to \( k=1 \).
3) Equate to zero, in turn, each non-zero weight, \( w_{ij}^{(k-1)} \), of \( \mathbf{W}^{(k-1)} \), and determine the optimal upper-layer weight vector, \( \mathbf{v} \), minimizing \( F=\sum (y-\hat{y})^2 \) by linear least squares. Save the corresponding value of the objective function, \( F_{\mathbf{v}}^{(k)} \).
4) Find the minimum of the objective function values, \( \mathbf{\tilde{v}}^{(k)}=\min \{ F_{\mathbf{v}}^{(k)} \} \). Set \( \mathbf{W}^{(k)}=\mathbf{W}^{(k-1)} \) and equate to zero the weight corresponding to the minimum objective function value, \( w_{ij}^{(k)}=0 \) with \( \mathbf{\tilde{v}}=\arg \min \{ F_{\mathbf{v}}^{(k)} \} \).
5) Set \( \mathbf{\Psi}^{(k)}=k \) and save this variable in a matrix, \( \mathbf{\Psi}=\{ \mathbf{\Psi}^{(k)} \} \) (with the same dimension as \( \mathbf{W} \)).
6) Set \( k=k+1 \). If \( k>m \cdot N \), go to 3. Else, end.

The book-keeping matrix, \( \mathbf{\Psi} \), stores the iteration number at which each connection weight has been deleted: By studying the elements in the columns (or rows) of the matrix it is afterwards easy to deduce when a certain input (or hidden node) has been eliminated. Usually, the model performance is reported as the mean error, \( e=\sqrt{F/n} \), where \( n \) is the number of observations.

The computational effort of the method can be reduced considerably by going through the weights in the order of the hidden nodes they refer to (i.e., starting with the connections between the inputs and the first hidden node, etc.). First, the net input to each hidden node, \( a_i(t)=\sum_{j=1}^{N} w_{ij}x_j(t), i=1, \ldots, m \), at each “time instant”, \( t \), is determined, as well as the corresponding output, \( z_i(t)=\sigma(a_i(t)) \). At step 3 of the algorithm a resetting of \( w_{ik} \) simply means that the net input of the \( i \)-th hidden node is changed into \( a_i(t)=-w_{ik}(t)x_k(t) \), while the net inputs, and outputs, of all other hidden nodes remain unaltered. Thus, for each weight in step 3, only one multiplication and one subtraction is needed to get the net input. In addition, the sigmoidal transformation is required for the hidden node in question. Along with the progress of the algorithm, the computational burden decreases gradually due to the permanent pruning of the weights, and, finally, of the hidden nodes.

A shortcoming of nonlinear black-box models is that they may yield poor predictions on independent data if some of the important inputs experience a level change. This is quite common in industrial processes, where the inputs may change rather dramatically because of control actions and changes in raw materials. A remedy follows logically from the over-all approach taken in the algorithm of this paper; it is natural to adjust the upper-layer weights as new information enters, and this is a linear problem, since a given input vector, \( \mathbf{x}(t) \), and a fixed set of lower-layer weights yield a fixed output, \( \mathbf{z}(t) \), from the hidden nodes for every time instant \( t \). For updating the upper-layer weights, \( \mathbf{v} \), the well known (recursive) Kalman filter \(^{14}\) can be used

\[
\mathbf{v}(t+1)=\mathbf{v}(t)+\mathbf{K}(t+1)(\mathbf{y}(t)-\hat{y}(t)) \quad \cdots \quad (1)
\]

where the Kalman gain is given by

\[
\mathbf{K}(t+1)=\frac{\mathbf{P}(t)\mathbf{z}^T(t+1)}{1+\mathbf{z}(t+1)\mathbf{P}(t)\mathbf{z}^T(t+1)} \quad \cdots \quad (2)
\]

while the matrix \( \mathbf{P} \) is updated by

\[
\mathbf{P}(t+1)=\mathbf{P}(t)+\mathbf{R}-\mathbf{K}(t+1)\mathbf{z}(t+1)\mathbf{P}(t) \quad \cdots \quad (3)
\]

In the equations, the covariance matrix of the “measurement error” is \( \mathbf{R} \), which, without a priori information, is usually chosen as a diagonal matrix, \( \mathbf{I} \). If \( c \) is large the gain \( \mathbf{K} \) is high, which makes large weight changes possible. The initial value of the matrix, \( \mathbf{P}(1) \), is usually chosen as a diagonal matrix with large diagonal elements. If the training set is directly followed by the test set, which is often the case in on-line learning, the results can be made less dependent on \( \mathbf{P}(1) \) by first applying the filter on the training set, using the final weights and matrices as starting points for the predictions.

3. Hot Metal Silicon Prediction

This section first describes the background of the hot metal silicon problem and briefly reviews earlier data-driven modeling efforts in the field, followed by a description in Subsec. 3.2 of the data set used for the analysis in the present work. Next, the overall behavior of the algorithm outlined in Sec. 2 is illustrated, with a more detailed description of a model candidate in Subsec. 3.4. Subsection 3.5 presents a statistical analysis of multiple runs with the algorithm (from different starting points), which results in a more accurate picture of the relevant input variables, and their time lags. The last subsection presents and evaluates a final model candidate formed on the basis of the findings.

3.1. Background

The blast furnace is the principal unit in the most important process route for iron produced for primary steelmaking. It is a complicated multi-phase process with heat and mass transfer, chemical reactions and phase changes, showing distributions of the state variables in both space and time. Large time delays of changes in the solid/molten phases and sluggish response to control actions make it important to predict quality and operational variables, e.g., the composition of the hot metal. The silicon content of the hot metal is important for the quality of the iron, but also as an indicator of the thermal state of the furnace. It reflects the energy available in (or vertical extent of) the high-temperature region, so changes in the “energy reserve” of the
process are usually seen as fluctuations in the hot metal silicon content. A decreasing silicon content often indicates a cooling of the furnace that without due countermeasures can lead to serious operational complications, while a high silicon content indicates excessive generation of heat, waste of coke, which also adversely affects the downstream processes. Blast furnaces are usually operated with a safety margin, i.e., a slightly higher coke rate than is deemed necessary. Since the cost of coke is dominating in ironmaking, there are obvious economical benefits of reducing the safety margin, but this requires stricter control of the heat level. Even though the main mechanisms behind the silicon transfer have been clarified, i.e., the gasification of SiO$_2$ from coke and coal ash to SiO(g) in the high temperature region at the tuyere level, and its reduction to silicon in the liquid iron, the most successful approaches to short-term prediction (e.g., on a tap to tap basis) of the hot metal silicon content are based on stochastic models because of uncertainty in the estimates of the internal state and dynamics of the high-temperature region. Numerous black-box models for the prediction of the hot metal silicon content have been developed, and some recent papers have stressed the chaotic nature of the signal. In most of these efforts, the input variables in the models have been selected on the basis of process knowledge, but in some of the papers more systematic approaches have been made. For instance, Waller and Saxén applied an exhaustive search among linear finite impulse response (FIR) models using a large set of inputs with different time lags, Bhattcharyan used a partial least squares procedure to select relevant inputs, while Saxén et al. evolved sparsely connected neural network models of the silicon content by a genetic algorithm. However, in practically all approaches on nonlinear prediction of the silicon content by neural networks, a small set of potential inputs was always selected a priori. The reason for this is, obviously, the considerable numerical effort required for training the networks.

### 3.2. The Data Set

The present analysis was based on a data set from a Nordic blast furnace, where the variables have been preprocessed to yield hourly mean values. The hot metal silicon content is not generally recorded regularly, but is measured for every batch (torpedo, ladle or mixer). However, in the furnace studied here the analysis was obtained with measurements obtained with a frequency of roughly 1 h$^{-1}$. On the basis of process knowledge, one-hour mean values of the 15 potential input variables reported in Table 1 were selected: Specific (sp.) quantities are expressed per ton of hot metal (thm). The table also presents the variables’ internal numbering in the paper, their units as well as the symbols that will be referred to later. Some of the variables are directly measured, while others are quantities readily computed. In order to evaluate the possibility of on-line prediction of the silicon content, the problem was written as

$$\hat{y}(t) = f(x(t), x(t-1), x(t-2), \ldots, x(t-8)), \ldots,$$ (4)

including lags of the 15-dimensional input vector, $x$, up to 8 h. The dimension of the input vector is thus $15 \times 9 = 135$. Note that autoregressive terms were deliberately omitted, since the inclusion of such is known to yield models of high inertia and with small possibilities to predict rapid changes in the output. Data for 800 h of operation was used for training, while the 187-hour period that follows was used for evaluating the performance of the resulting models. Before the models were developed, each (input and output) variable was scaled to the interval (0,1) by dividing the difference between the original value and the minimum value by the range of the variable in the training set.

### 3.3. Overall Behavior

The performance of the algorithm is first illustrated by a run of it from a random starting weight matrix, using a network with five hidden nodes. (An analysis with a higher number of hidden nodes illustrated that similar final performance of the algorithm was obtained.) The Kalman filter was initialized by $P(1) = 10I$ (at the start of the training set) while $R=I$. Figure 2 illustrates the evolution of the root mean square errors, $\epsilon$, on the training set (solid lines), test set without (dashed lines) and with (dotted lines) online adaptation of the upper-layer weights. The algorithm, which progresses from right to left in the figure, is seen to initially reduce both the training and test errors (without online learning) considerably, but the latter ones remain on a considerably higher level throughout the pruning. By contrast, the continuously updated model only shows a slowly decreasing trend, but throughout the run it yields errors comparable with the training errors. Thus, the weight updating results in a model that generalizes well.

From the run the lowest error for the on-line-updated version, $\epsilon=0.073$, is encountered when only six connections remain in the lower layer of the network. Figure 3 illustrates the approximation provided by this network (dotted lines), the target of the test set (solid line), and the results with fixed upper layer weights (dashed lines). The online learning is seen to lead to clearly superior performance: It has not only removed the bias in the prediction but has also been able follow the large changes in the signal.

### Table 1. Potential input variables in the hot metal silicon content modeling.

<table>
<thead>
<tr>
<th>Number</th>
<th>Variable</th>
<th>Unit</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Total blast volume</td>
<td>m$^3$/h</td>
<td>$V_{bt}$</td>
</tr>
<tr>
<td>2</td>
<td>Blast pressure</td>
<td>bar</td>
<td>$p_{bl}$</td>
</tr>
<tr>
<td>3</td>
<td>Gas permeability</td>
<td></td>
<td>$\kappa$</td>
</tr>
<tr>
<td>4</td>
<td>Sp. coal injection</td>
<td>kg/thm</td>
<td>$m_{c}$</td>
</tr>
<tr>
<td>5</td>
<td>O$_2$ content of blast</td>
<td>%</td>
<td>$X_{CO}$</td>
</tr>
<tr>
<td>6</td>
<td>Gas CO utilization</td>
<td>%</td>
<td>$\eta_{CO}$</td>
</tr>
<tr>
<td>7</td>
<td>Top gas CO+CO$_2$</td>
<td>%</td>
<td>$r_{CO+CO2}$</td>
</tr>
<tr>
<td>8</td>
<td>Flame temperature</td>
<td>°C</td>
<td>$T_8$</td>
</tr>
<tr>
<td>9</td>
<td>Coke rate</td>
<td>t/h</td>
<td>$m_{c}$</td>
</tr>
<tr>
<td>10</td>
<td>Coal rate</td>
<td>t/h</td>
<td>$m_{c}$</td>
</tr>
<tr>
<td>11</td>
<td>Energy at tuyeres</td>
<td>MW</td>
<td>$E_{ty}$</td>
</tr>
<tr>
<td>12</td>
<td>Sp. blast volume</td>
<td>m$^3$/thm</td>
<td>$V_{bt}$</td>
</tr>
<tr>
<td>13</td>
<td>Solution loss rate</td>
<td>kg/thm</td>
<td>$m_{c}$</td>
</tr>
<tr>
<td>14</td>
<td>Tuyere heat loss</td>
<td>MW</td>
<td>$Q_{ty}$</td>
</tr>
<tr>
<td>15</td>
<td>Ore/Coke ratio</td>
<td></td>
<td>$\phi$</td>
</tr>
</tbody>
</table>

© 2007 ISIJ 1734
3.4. Analysis of a Model Candidate

An analysis of this network reveals that the relevant input variables are #3, #5, #7, #12 and #14, i.e., the gas permeability, the oxygen content of the blast, the sum of the top gas CO and CO2, the specific blast volume as well as the tuyere heat loss: The last variable, which contributes by two inputs (at different lags), is the heat flow (in kW) from the tuyeres to the cooling water, calculated on the basis of water flow rate and temperature measurements. In terms of relevance of the variables, with the most important one (i.e., last remaining input) listed first among the arguments, the model can be written as

\[ \hat{\text{Si}}(t) = f(Q_{\text{tuy}}(t-2), V_{\text{bl}}(t-1), k(t-2), Y_{\text{CO}_2}(t-3), Q_{\text{tuy}}(t-3), X_{O_2}(t-4)) \] ......... (5)

Thus, the heat loss at the tuyeres (lagged by 2 and 3 h), the specific blast volume (lagged by 1 h), the gas permeability factor (lagged by 2 h), the top gas CO+CO2 content (lagged by 3 h) and the blast oxygen content (lagged by 4 h) are considered to be relevant for predicting the silicon content. These variables are mostly related to the tuyere parameters, which are known to affect the silicon transfer in the furnace. The selected variables, as well as their comparatively short time lags (maximum 4 h), are findings that agree well with knowledge from the practical operation of the furnace: The residence time of the molten iron between the tuyere level and tapping is in the order of 2 h. The tuyere heat loss is known to reflect the intensity of the thermal conditions in the lower furnace, the specific blast consumption is a measure of the efficiency of the operation of the BF, while the gas permeability reflects the vertical extent of the high-temperature region.

This network has been depicted in Fig. 4, which shows that three hidden nodes are used: One for the specific blast volume, one for the oxygen content of the blast, and one for the two tuyere heat loss terms together with the gas permeability and the top gas CO+CO2 content. The weights reported to the left of the corresponding connections reveal that the contribution of the last hidden node is minor. This observation is also supported by the fact that the next step of the pruning algorithm eliminates the oxygen content from the input set. Another remark is that all inputs, except the sum of CO and CO2 in the top gas, exhibit positive correlation with the silicon content. The only negative correlation is understood in that \( Y_{\text{CO}_2} \) describes the (major) outflow of carbon from the process, and a sudden increase in this means that the fuel reserved is depleted, resulting in an internal cooling and a lower silicon content.24)

3.5. Detection of Relevant Inputs and Time Lags

Since a run of the algorithm on this specific problem with about one hundred inputs only takes about 30 s in MATLAB® —a numerical computing environment and programming language provided by The MathWorks25)—on a standard PC, it is clearly feasible to carry out repeated runs from different starting points. In order to study more general features of the problem, a larger number of runs with the algorithm were undertaken, starting from 30 random initial weight matrices. The most relevant ten variables, i.e., the ones that remained unpruned until the last ten steps, were recorded. Figure 5 presents the occurrence, \( P \), of the 135 input signals in this set: Each of the 15 different input variables (cf. Table 1) has a segment separated by ver-
tical dashed lines, within which the order goes from the largest to the smallest time lag. Thus, the first input signal (#1) is the blast volume lagged by 8 h, $V_{bl}(t/8)$, while the last signal (#15) is the unlagged value of the ore-to-coke ratio, $(o/c)(t)$. From this figure some interesting observations can be made: Some of the inputs (e.g., flame temperature (#8), top gas CO utilization (#6) and coke solution-loss rate (#13)) are seen to be of minor importance, while some (e.g., specific (#4) and total (#10) coal injection rates) show correlation with the output but without a preference for a specific time lag. As for the most relevant variables, the tuyere heat loss, $Q_{tuy}(t/2)$, is clearly dominating, followed by the gas permeability, $\kappa$, (#3) and the specific blast volume, $V_{bl}/H(\#12)$. All three variables follow the same pattern further illuminated in Fig. 6; they exhibit a pronounced peak at a time lag of 2 h, but also a clear autocorrelation. For the former two, it is striking to note that their unlagged values (i.e., at time $t$) are considered irrelevant—a fact that can be justified by the residence time of the hot metal in the hearth. Focusing on the most relevant variable only, i.e., the last remaining one at the end of the pruning process, it was found that $Q_{tuy}(t-2)$ occurred 12 times, followed by $V_{bl}(t-2)$ and $V_{bl}(t-3)$ (both with three occurrences). Also this shows the strong correlation between the silicon content and the tuyere heat loss and the specific blast volume.

These findings also show agreement with the results of the specific network presented in Fig. 4, and also with the general findings of earlier efforts to predict the hot metal silicon content in other blast furnaces. For instance, the heat loss,16,11) and gas permeability 11) were found important. In an exhaustive search 12) the most important inputs found were mainly related to the gas permeability and the energy and heat loss at the tuyeres. The recent study on evolving neural network models by a genetic algorithms23) also found the tuyere heat loss, the specific blast volume and the gas permeability important. Also other investigators20) have found permeability indices relevant for predicting the hot metal silicon content.

A more detailed statistical analysis of the most promising models was undertaken in order to find a proper lower-layer connectivity of the network in a final model of the silicon content. The connections in all of the last five networks of the 30 runs were analyzed, yielding the occurrences reported in Table 2. Quite naturally, the most frequently occurring connections to the hidden nodes are those from a single input, and the top four inputs reported earlier also reappear here, but in a slightly different order: The tuyere heat loss, lagged by 3 h, occurs more frequently than the specific blast volume, lagged by 2 h or 3 h. Also the gas permeability (lagged by 2 h) is a frequent variable and occurs more than ten times, but the remaining variables with a single connection appear clearly less often (8 times). As for the most frequent connections from two inputs to a hidden node, the occurrences are, obviously, much fewer. Again, the relevant variables detected above are strongly represented (cf. Table 2).

### 3.6. Development of the Final Model

On the basis of the results presented in the preceding...
subsection, the compromise model illustrated in Fig. 7 is suggested. It is motivated by the findings presented in Table 2 in that the single input with the highest occurrence, $Q_{\text{net}}(t-2)$, is the only input to one hidden node, the most important combination of two variables, $x(t-2)$, $V_{\text{col}}(t-3)$, occurs as input to another hidden node, and, finally, the second most important single variable, $Q_{\text{net}}(t-3)$, is combined with the specific blast volume at time lag three, $V_{\text{col}}(t-3)$, to a final hidden node. The latter combination also occurs high on the list of input pairs (with $P=3$). This choice of inputs reduces the number of hidden nodes, yielding only twelve parameters (including the biases) in the network. The model is also appealing in that it predicts the silicon content 2 h ahead.

Training of all parameters of this network to completion using an efficient gradient based method for neural network training resulted in training and test errors of 0.093 and 0.131 respectively, while the on-line updated model showed an error of $E=0.074$, i.e., of same accuracy as the best models evolved by the pruning process, as illustrated in Fig. 8. The latter model has explained 55% of the variation in the silicon content. This demonstrates that a very parsimonious model can be developed on data from a Nordic blast furnace demonstrated that the inputs found are variables that are known to correlate with the hot metal silicon content. Furthermore, the detected time lags are reasonable considering the dynamics of the process. The merits of an on-line updating of the upper-layer weights for practical application of the models developed were further demonstrated. In a final attempt, the relevant inputs and model structure were analyzed by a full gradient-based training of the weights of existing (sparse) connections. This was demonstrated to be fast, robust and straightforward to use without a need to set model-specific parameters. A statistical analysis of the results of several runs of the algorithm on data from a Nordic blast furnace demonstrated that the inputs found are variables that are known to correlate with the hot metal silicon content. This demonstrated to yield a final model that was able to carry out accurate one-step ahead predictions of the silicon content.

4. Conclusions

This paper has outlined an algorithm for the selection of input variables, their time lags as well as a proper complexity of a multi-layer feedforward neural network in a time-series prediction task, and has described an application of the technique to the prediction of the silicon content of hot metal produced in a blast furnace. The algorithm has been demonstrated to be fast, robust and straightforward to use without a need to set model-specific parameters. A statistical analysis of the results of several runs of the algorithm on data from a Nordic blast furnace demonstrated that the inputs found are variables that are known to correlate with the hot metal silicon content. Furthermore, the detected time lags are reasonable considering the dynamics of the process. The merits of an on-line updating of the upper-layer weights for practical application of the models developed were further demonstrated. In a final attempt, the relevant inputs and model structure were analyzed by a full gradient-based training of the weights of existing (sparse) connections. This was demonstrated to yield a final model that was able to carry out accurate one-step ahead predictions of the silicon content.

REFERENCES