In the past half century, developments in control theory have been successfully applied to the lower level control of industrial processes [1]-[5]. In contrast with conventional machine control, however, industrial process control requirements concern product performance instead of the accuracy of the lower level control loop. The process performance is an overall index, including technical and economic specifications, which are often difficult for the lower level control system to handle directly and solely. Higher level supervision is often required to adjust the process control system from time to time. Since many industrial processes are of a complex nature (e.g., highly nonlinear, seriously coupled, higher order, and time varying), it is difficult to develop a closed-loop control model for this higher level supervision [6], [7], even though some advances have been achieved in hybrid control design theory [8]. Thus, the human operator is often required to provide online adjustment, which makes the process performance greatly dependent on the experience of the individual operator. Furthermore, process control has become a case-by-case phenomenon [9], [10] and has less referential value. It would be extremely useful if some kind of systematic methodology can be developed for the process control model that is suited to one kind of industrial process.

Traditionally, there are two ways to develop process control methods. The first is the knowledge-driven (white-box) method. In this method, mechanistic models are derived from physico-chemical laws, which can include qualitative information in the form of expert and/or linguistic knowledge. This kind of method is well suited to a wide range of process operations [11] and is still being widely used in industry, but its case-by-case characteristics can make its application difficult and time consuming for industrially relevant control systems. The second is the data-driven (black-box) method motivated by development of identification and intelligent techniques. Linear and nonlinear ARMAX-type models are widely used to model dynamic systems [12], and intelligent methods such as fuzzy logic [13], neural networks [14], genetic algorithms [15], rule-based expert systems [16], and even multivariate statistical partial least squares [17] have been proposed to model the dynamics of industrial processes. However, this type of data-driven model requires extensive data or experience and may not be suited for newly...
developed processes where data and knowledge are scarce. Furthermore, its reliability can be suspect due to its incapability to predict behavior outside the experimental datum domain.

Recently, gray-box approaches have been developed [18], [19] that combine the two methods. Using the knowledge-driven method, the model can have a relatively simple structure; using the data-driven method, its parameters can be estimated and its learning ability can be improved. Although this kind of approach appears suitable for the control of complex industrial processes, it still requires a systematic effort to select, organize, and coordinate many different methods into one efficient, practical framework for process control.

Motivated by the above problems, a hybrid control methodology is proposed in this article for supervisory control of distributed control system (DCS)-based batch processes. This hybrid intelligent control methodology combines primary and auxiliary control structures, knowledge- and data-driven methods, and conventional and intelligent techniques into one efficient control model. Although each control element is well known, their innovative combination, which consists of nominal and supplementary parts, can generate better and more reliable performance. Based on the knowledge-driven method, the primary controller provides proper set-points for the DCS to achieve nominal performance. Based on the hybrid knowledge/data-driven method, the auxiliary compensator provides corrections to DCS set-points according to the dynamic environment. Thus, process performance will be improved from the nominal condition instead of from scratch. On the other hand, any failure of the compensation system will have little effect on the original performance. This hybrid strategy greatly improves both the performance and reliability of the process. Practical experiments on a DCS-controlled laminar cooling process in the metallurgical industry show the effectiveness and promise of the methodology for use in the manufacturing industry.

**DCS-Controlled Laminar Cooling Process**

**System Description**

The laminar cooling process in the metallurgical industry, simplified in Fig. 1, is controlled by the DCS to cool a metal slab from finishing temperatures (820-900 °C) down to the straightening temperatures (500-700 °C) according to the slab thickness and the steel grade. Poor cooling will cause an uneven temperature drop throughout the slab, resulting in quality deterioration. More seriously, the deformed slab may damage the processing machine and abort the whole operation. A poorly controlled laminar cooling process may cause significant economic loss to both manufacturer and customer.

Four infrared pyrometers (\(T_i - T_j\)) are mounted on top of the runout table to measure the top surface temperature of the slab at exit from the mill, at entry to the cooling area, at exit from the cooling area, and at entry to the straightener. Two pyrometers (\(T_i^*\) and \(T_j^*\)) are mounted on the bottom of the runout table to measure the bottom surface temperature of the slab. One X-ray gauge (\(D_i\)) is located at the exit of the rolling mill. The cooling area is partitioned into three sections, labeled \(A\), \(B\), and \(C\). Thirteen cooling units are uniformly spaced along section \(B\). The number of cooling units in operation, \(N\), the water flow of one top cooling unit, \(Q_r\), and the water flow rate, \(\gamma\) (bottom flow/top flow), can be adjusted separately. The temperature drop is caused by the heat radiation only in sections \(A\) and \(C\) and both radiation and water cooling in section \(B\) [20].

The technical targets of the cooling process refer to:

- the temperature loss rate \(V_r\) caused by water cooling;
- the final cooling temperature error \(\Delta T_{cm}(=T_i - T_{cm})\), where \(T_{cm}\) is the final cooling temperature measured by pyrometer \(T_i\) and \(T_j\) is the desired final cooling temperature;
- the temperature deviation \(\Delta T_{bm}\) between the bottom (lower) and top (upper) surfaces of the slab at pyrometers \(T_i^*\) and \(T_j^*\);
- the temperature difference \(\Delta T_{bm}\) between the head and end of the slab at pyrometer \(T_i^*\).

**Operating Modes and Operating Points**

In the cooling process, the operating mode is defined by the steel grade, the thickness \(d\), and the final cooling temperature \(T_{cm}\). In any operating mode, the main entry boundary conditions for the slab to move into the cooling area are the thickness \(d\), detected by X-ray gauge \(D_i\), and the temperature \(T_{cm}\), measured by the heat pyrometer \(T_r\). These two boundary condition variations (\(\Delta T_{cm}\) and \(\Delta d\)) can be regarded as measurable disturbances and have much influence on the cooling process. Other boundary conditions, together with other internal changes in the cooling process, will be treated as small disturbances and expressed as an unobservable vector \(\Omega\), which includes:
- the temperature deviation \(\Delta T_{bm}\) between the bottom and top surfaces of the slab \(i\) before pyrometer \(T_i^*\);
- the temperature deviation \(\Delta T_{bm}\) between the bottom and top surfaces of the slab \(j\) before pyrometer \(T_j^*\);
- the temperature deviation \(\Delta T_{bm}\) between the bottom and top surfaces of the slab \(k\) before pyrometer \(T_k^*\);
- the temperature deviation \(\Delta T_{bm}\) between the bottom and top surfaces of the slab \(l\) before pyrometer \(T_l^*\).

![Schematic diagram of the laminar cooling system. \(T_i - T_j\): four top pyrometers; \(T_i^*\), \(T_j^*\): two bottom pyrometers; and \(D_i\): X-ray gauge.](image-url)
• the temperature difference $\Delta T_{in}$ between the head and end of the slab $i$ before pyrometer $T'_d$;
• the temperature variance of water $\Delta T_{sw}$ and the environment $\Delta T_{en}$.

The operating point consists of three different types of variables:

- Index variables: steel grade SG, thickness $d_i$, and desired cooling temperature $T_{cm0}$. They are regarded as the indexes for searching an appropriate operating point in the database, and some models also employ them occasionally.
- Boundary variable: slab entry temperature $T_{cm0}$, which is the boundary condition for the model computation.
- Control variables: acceleration $a$, flow rate $\gamma$, speed $v$, cooling units $N$, and water flow $Q_e$, which are also set-points of the DCS. The subscript zero refers to a nominal value.

Generalization of a Human-Supervised, DCS-Based Process

The laminar cooling process can be described approximately by two subprocesses: the uncontrollable model $P_1$ and the complex model $P_2$ (see Fig. 2). The measurable disturbances $\Delta T_m$ and $\Delta d$ influence the cooling process through model $P_1$. The unmeasurable disturbances $\Omega_i$ affect the cooling process through model $P_2$.

In the laminar cooling process, the DCS provides four separate control loops: 1) a water flow ($Q_e$) control loop; 2) a flow rate ($\gamma$) control loop; 3) a runout table speed/acceleration ($v/a$) control loop; and 4) a cooling units ($N$) control loop. What the DCS can handle at the lower level are control variables $v_i$, $a_i$, $\gamma_i$, $N_i$, and $Q_{e,i}$. These local control loops are well stabilized by the DCS. Performance requirements are in terms of temperature variables $T_{cm0}$, $T_{sw0}$, and $T_{en0}$, which can only be maintained via proper values of the above control variables. A hidden dynamic exists between what the DCS can directly control and the process performance. This unknown dynamic is affected by many elements, such as the slab temperature, the thickness and shape of the slab, the slab material, the temperature of the water and the environment, etc.; it is impossible to identify all of them [21], [22]. Therefore, it will be difficult for the DCS to provide proper set-points at a higher level for maintaining satisfactory performance.

To maintain satisfactory performance, human experts are often required, as shown in Fig. 2, to determine the proper set-point values for control variables through several typical experiments and a period of manufacturing operation. The supervisory control objectives can be summarized as follows:

1) To obtain and process the data from several typical experiments;
2) To establish the adjustment relation between the control variables and the controlled errors, and the internal relations between control variables, according to previous experience;
3) To determine a proper sequence for regulation.

Control Actions

At each operating mode, the loss rate $V_r$ is relatively constant and defined by the technical requirement

$$ V_r = \frac{\Delta T}{\Delta t} = \frac{\Delta T_{sw}}{\Delta L} $$

(1)

where $\Delta T = \text{the temperature drop (°C) caused by the cooling water; } \Delta t = \text{the time period over which the slab stays in the cooling area; } \Delta L = \text{the length of the cooling area; and } v = \text{the speed of the runout table. In general, } \Delta T/\Delta L \text{ is relatively constant, so speed } v \text{ is the main factor in determining } V_r. \text{ Since the speed of the runout table } v \text{ can be defined prior to each cooling mode, } V_r \text{ can be determined offline instead of online.}$

Through physical analysis, the higher level supervisory control can be determined as follows:

- Reduce $\Delta T_{sw}$ by accelerating the runout table $a$.
- Eliminate $\Delta T_{sw}$ by adjusting the flow rate $\gamma$.
- Control $T_{cm0}$ or $T_{en0}$ by the number of opened cooling units $N$ and the water flow $Q_e$.

However, the strong nonlinear coupling between these control variables makes the above supervision difficult.

Hybrid Intelligent Control Methodology

If the boundary conditions can be maintained to be invariant (i.e., disturbances such as $\Delta T_{in}$ and $\Delta d$ can be eliminated), the proper operating point will be found after a small number of slab cooling trials. Unfortunately, in a manufacturing site, many disturbances (measurable and unmeasurable) exist and make it difficult for the operator to obtain the proper operating point. Since the experience and knowledge of the human operator provide a good qualitative model.
for the supervisory system, intelligent technologies are a good candidate for this unmodeled process. Since the technical indexes $T_{cm}$, $\Delta T_{du}$, and $\Delta T_{he}$, upon which the human operator bases a decision, are statistical values, a statistical process controller (SPC) is necessary to generate them.

Based on the DCS-controlled laminar cooling process in Fig. 2, a framework for the hybrid supervisory control system, consisting of a primary controller and an auxiliary compensator, is established as shown in Fig. 3. The primary controller is intended to replace the human operator and the auxiliary compensator to enhance performance. The major components of the system are described below.

1) An SPC system is used to simulate the human perception of input signals.

2) A primary control system provides the proper operating points of control variables for each operation mode:
   - A database is used to store the operating points of all the operating modes;
   - An autosearch mechanism is used to explore the new operating point for the required operating mode.

3) An auxiliary compensator consisting of three different compensation models (coarse, general, and fine) is used to compensate the effects of the variation in boundary conditions according to the operating environment.

The operating phases of supervision, which consist of an initialization phase and three operational phases, are summarized in Table 1. When the new slab comes into the cooling process, the primary controller will check the operating point from the database with the signals $SG$, $d$, and $T_c$. If there is no operating point available for the new slab, a proper operating point will be determined via the autosearch mechanism of the primary controller in Fig. 4 through several experiments. The operating point variables provided by the primary controller are considered as nominal values. Once the proper operating point is determined, the primary controller will be switched off. The nominal values of DCS set-points are provided by the primary controller and finely tuned by the auxiliary compensator according to operating conditions.

### Statistical Process Control

Since the process performance is evaluated via the human operator using the statistical variables $T_{cm}$, $\Delta T_{du}$, and $\Delta T_{he}$, an SPC is required to process the temperature outputs at pyrometers $T_4$, $T_3$, and $T_3'$. The SPC actually performs the human perception function in Fig. 2. The outputs of this function are defined as

$$T_{cm} = \sum_{i=1}^{K} T_{4(i)} / K$$

$$\Delta T_{du} = 2 \left( \sum_{i=1}^{K/2} T_{4(i)} - \sum_{i=K/2 + 1}^{K} T_{4(i)} \right) / K$$

$$\Delta T_{he} = \sum_{i=1}^{K} (T_{3(i)} - T_{3'(i)}) / K$$

(2)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Supervision</th>
<th>Primary Controller</th>
<th>Auxiliary compensator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>on</td>
<td>on</td>
</tr>
<tr>
<td></td>
<td>Initial search for proper set-points</td>
<td>on</td>
<td>on</td>
</tr>
<tr>
<td>2.1</td>
<td>Operation with coarse compensation</td>
<td>off</td>
<td>on</td>
</tr>
<tr>
<td>2.2</td>
<td>Operation with general compensation</td>
<td>off</td>
<td>off</td>
</tr>
<tr>
<td>2.3</td>
<td>Operation with fine compensation</td>
<td>off</td>
<td>off</td>
</tr>
</tbody>
</table>

June 2001 IEEE Control Systems Magazine
where $K$ is the number of samples, which depends on the sampling interval and the time for the slab to move over pyrometers $T_1$, $T_2$, and $T_3$.

**Fuzzy Autosearch System**

As shown in Fig. 2, the human operator adjusts the set-point variables of the DCS ($\gamma$, $a$, $N$, $Q_e$) via his or her perception of process variables $\Delta T_{in}$, $\Delta T_{in}$, and $\gamma$. A multivariable fuzzy system can be used to simulate the decision making of the human operator. This high-dimensional fuzzy system is difficult to construct due to couplings between control loops. Qualitatively, the design of the multivariable fuzzy control system can be simplified via the following two steps:

- Design a conventional low-dimensional fuzzy controller for each dominant loop;
- Design a qualitative decoupling mechanism to suppress the coupling between loops.

Through theoretical analysis and extensive simulations, this multivariable process can be considered as four dominant loops with couplings between the set-point variables, as shown in Table 2.

### Decomposition of Multivariable Fuzzy Control System

Four dominant fuzzy logic controllers ($F_1$, $F_2$, $F_3$, and $F_0$) are used to control four set-point variables ($\gamma$, $a$, $N$, $Q_e$) according to input conditions $d$, $\Delta T_{in}$, $\Delta T_{in}$, and $\gamma$. When one loop is operating in the isolated situation (other loops are not in operation), the interaction from other loops is minimal. The fuzzy inference system is thus relatively easy to design. Since the fuzzy inference system is used to search for proper DCS set-points, a PI-type fuzzy controller [23] is suitable for eliminating the steady-state error. The membership functions can be chosen as triangular, and rules should be extracted from fundamental knowledge and human experience about the process.

### Qualitative Decoupling Mechanism

Through fundamental analysis and extensive experimentation, the coupling between control loops and the basic human strategy to address it are summarized below.

1) Classification of coupling
   a) Both $N$ and $Q_e$ are for adjustment of $T_{in}$ only;
   - Variable $N$ is for large and coarse adjustment and $Q_e$ for fine adjustment.
   b) Variables $\gamma$ and $a$ affect all the performance indexes: $\Delta T_{in}$, $\Delta T_{in}$, and $T_{in}$.
   - Variables $\gamma$ and $a$ have more influence on $N$ and $Q_e$ because variations of $\gamma$ and $a$ will not only change $\Delta T_{in}$ and $\Delta T_{in}$ effectively, but also greatly change $T_{in}$.
   - Variations of $N$ and $Q_e$ have little influence on $\gamma$ and $a$ because variation of $T_{in}$ has little effect on $\Delta T_{in}$ and $\Delta T_{in}$.

2) Human strategy to address coupling
   - First, the variables $\gamma$ and $a$ should be tuned to eliminate $\Delta T_{in}$ and $\Delta T_{in}$.
   - After obtaining satisfactory $\Delta T_{in}$ and $\Delta T_{in}$, the final cooling temperature $T_{in}$ can be controlled by adjusting $N$ and $Q_e$ on which $\Delta T_{in}$ and $\Delta T_{in}$ have little effect.

3) Fuzzy decoupling system

The above human decoupling strategy can easily be developed using the following three rules:

- Coarse adjustment: If $\Delta \gamma$ or $\Delta a$ are relatively large (PL, PB, PM, PMB, PM, NM, NMB, NB, NL), or $\Delta T_{in}$ is large (PL, PB, NB, NL), then start the inference $F_0$ to adjust $N$.
- Fine adjustment: If $\Delta \gamma$ and $\Delta a$ are relatively small (PMS, PS, PZ, NS, NS, NMS) and $\Delta N$ is zero, then start the inference $F_0$ to adjust $Q_e$.
- Regular adjustment: Stop inference $F_0$ and $F_0$ in other circumstances.

In these rules, the codes NS, PZ, PM, PB, PL, etc., are commonly used acronyms for fuzzy sets. N stands for negative, P stands for positive, Z stands for zero, S stands for small, M stands for medium, B stands for big, and L stands for larger than big.

Due to this decoupling mechanism, a high-dimensional multivariable fuzzy system becomes feasible by using a number of low-dimensional fuzzy rule bases. Since the fuzzy rules are normally extracted according to one type of steel grade (e.g., low-carbon steel, 0.23%C), an adaptive factor $\eta$ is used to adjust the inference results for different types of steel:

$$\eta = k_0 / k$$

### Table 2. Dominant relationship between the DCS set-points and process outputs.

<table>
<thead>
<tr>
<th>Set-point adjustment</th>
<th>Process outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>$\Delta T_{in}$</td>
</tr>
<tr>
<td>$a$</td>
<td>$\Delta T_{in}$</td>
</tr>
<tr>
<td>$N$ and/or $Q_e$</td>
<td>$\Delta T_{in}$</td>
</tr>
</tbody>
</table>
where \( k \) is the thermal diffusivity of any steel grade and \( k_0 \) is the grade upon which the fuzzy rules are designed.

**Auxiliary Compensation System**

The auxiliary compensator eliminates the boundary condition perturbations. At the beginning stage, when data are scarce, a knowledge-based approach is used for coarse compensation. Later, when more data have been collected, a hybrid knowledge/data-based approach is suggested for general compensation with better accuracy. At the final stage, when extensive data have been accumulated for each operating mode, a data-based approach could be used for fine compensation, if necessary, to meet the highest adaptation and control accuracy requirements.

**Coarse Compensation**

For a new process where data are scarce, the knowledge-based approach may be the only option. Since adjustment of \( N \) and \( Q_e \) has fewer coupling effects on \( \gamma \) and \( a \), a mechanistic model for tuning \( N \) and \( Q_e \) can be developed, based on physico-chemical analysis and experiment [24]-[26], as follows:

\[
\Delta N_0 = \alpha_c \frac{T_{m0} d + d_a T_{m} + \Delta T_{m} d}{Q_{e0}} N_0 \tag{4}
\]

\[
\Delta N = \text{INT}(\Delta N_0 + 0.5) \tag{5}
\]

\[
\Delta Q_e = \left[\frac{N_0 + \Delta N_0}{N_0 + \Delta N} - 1 \right] \times Q_{e0} \tag{6}
\]

where \( \alpha_c \) is a constant coefficient, which can be modified by the human operator according to the difference in steel grade, and INT is the integer (truncation) function. The mechanistic model of (4)-(6) can suppress two major disturbance effects, \( \Delta T_m \) and \( \Delta d \) (variance between the current slab and the operating point).

\[
\Delta T_m = T_m - T_{m0} \quad \text{and} \quad \Delta d = d - d_0.
\]

Here \( T_{m0} \) and \( d_0 \) refer to the entry temperature and thickness at the operating point, and \( T_m \) and \( d \) refer to the entry temperature and thickness of the current slab. Other small disturbances, \( \Omega_s \), are neglected.

Although disturbances \( \Delta T_m \) and \( \Delta d \) influence flow rate \( \gamma \) and acceleration \( a \), the relations between them are difficult to describe from physico-chemical laws. Analytical adjustment of \( \gamma \) and \( a \) is not considered in the current coarse compensation.

**General Compensation**

After some data have been collected, the knowledge-based model can be enhanced with experimental data. A hybrid knowledge/data approach is proposed using an approximate linear model with self-adaptive ability. Because of the adaptive capability, better approximation can be expected with the hybrid approach.

**Approximate Linear Model (Knowledge Based):** According to the coupling classification in the previous section, the effects of \( \Delta N \) and \( \Delta Q_e \) on control variables \( \Delta \gamma \) and \( \Delta a \) can be omitted. Variables \( \Delta \gamma \) and \( \Delta a \) can then be approximated as linear functions of \( \Delta T_m \) and \( \Delta d \), as shown below, with coefficients \( \beta_{ij} \) to be identified:

\[
\Delta a = \beta_{11} \Delta T_m + \beta_{12} \Delta d \tag{7}
\]

\[
\Delta \gamma = \beta_{21} \Delta T_m + \beta_{22} \Delta d. \tag{8}
\]

This hybrid method combines primary and auxiliary control structures, knowledge and data-driven methods, and conventional and intelligent techniques into one efficient control model.

![Figure 5. Structure of the general compensation model.](image-url)
The model of $\Delta N$ and $\Delta Q_e$ in (7) and (8) can be further approximated as below with the detailed derivation given in Appendix 1:

$$
\Delta N = \text{INT}(\Delta N_1 + \Delta N_2 + 0.5) = \text{INT}\left( \frac{\Delta T_{cm} - T_0}{d_0} - N_0 + \Delta d N_0 + 0.5 \right)
$$

(9)

$$
\Delta Q_e = \beta_{31}\Delta N_1 + \beta_{32}\Delta N_2 + \beta_{33} \frac{\Delta N_1 + \Delta N_2 - \Delta N}{N_0 + \Delta N} Q_{e0} + \beta_{34} \Delta T_{cm} + \beta_{35} \Delta d.
$$

(10)

Equations (7), (8), (9), and (10) comprise the approximate linear model of the laminar cooling process. Unknown parameters $\beta_{ij}$ in (7), (8), and (10) need to be determined. Equations (7) and (8) can be transformed into a matrix formulation (11), and (10) can be transformed into (12)

$$
Y_i = \varphi_i^T \theta_i
$$

(11)

where

$$
Y_i = [\Delta \alpha \quad \Delta \gamma]^T
$$

$$
\varphi_i^T = [\Delta T_{cm} \quad \Delta d]
$$

$$
\theta_i = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}
$$

and

$$
Y_2 = \Delta Q_e
$$

$$
\varphi_2^T = \left[ \begin{array}{c} \Delta N_1 \\ \Delta N_2 \\ \Delta N_1 + \Delta N_2 - \Delta N \\ Q_{e0} \\ \Delta T_{cm} \\ \Delta d \end{array} \right]
$$

$$
\theta_2 = \begin{bmatrix} \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} & \beta_{35} \end{bmatrix}^T
$$

Modified Recursive Least-Squares (RLS) Identification (Data Based): Both (11) and (12) can be identified using the modified weighted RLS algorithm, which is the traditional weighted RLS [27] combined with an additional correction to the linear model:

$$
\begin{align*}
\dot{Y}_i^{(k+1)} &= Y_i^{(k+1)} - \Delta Y_i^{(k+1)} \\
K_i^{(k+1)} &= P_i^{(k)} \varphi_i^{(k)} [\lambda + \varphi_i^{(k)} P_i^{(k)} \varphi_i^{(k)}]^{-1} \\
\dot{\theta}_i^{(k+1)} &= \dot{\theta}_i^{(k)} + K_i^{(k+1)} [Y_i^{(k+1)} - \varphi_i^{(k)} \dot{\theta}_i^{(k)}] \\
P_i^{(k+1)} &= \frac{1}{\lambda_i} [I - K_i^{(k+1)} \varphi_i^{(k)}] P_i^{(k)}
\end{align*}
$$

(13)

where $i = 1,2$; $Y_i^{(k+1)}$ refers to control variables measured at the current sample time $k+1$; $\Delta Y_i^{(k+1)} = [\Delta \alpha, \Delta \gamma]^T$ and $\Delta Y_2^{(k+1)} = \Delta Q_e$ are corrections to control variables; $Y_i^{(k+1)}$ are updated control variables that should be used to control the process; $P_i^{(k+1)}$ and $K_i^{(k+1)}$ are intermediate variables for recursive computation; and $\lambda_i$ is the forgetting factor ($0 < \lambda_i < 1$).

Fuzzy Adaptation: Since the laminar cooling process model is unknown, it is difficult to obtain the correction $\Delta Y_i^{(k+1)}$. Practically, the fuzzy inference mechanism can be designed to roughly estimate the correction $\Delta Y_i^{(k+1)}$ from the boundary variable $d$ and the process errors $\Delta T_{cm}$, $\Delta T_{he}$, and $\Delta T_{in}$. Thus, the approximate linear control model can be adapted online as shown in Fig. 5.

A four-input/three-output fuzzy inference mechanism needs to be constructed, as well as the decoupling mechanism. Similar to the fuzzy auto-search system in Fig. 4, three low-dimensional fuzzy inference systems $F_1$, $F_2$, and $F_3$ are designed for correction of control variables $\gamma$, $a$, and $Q_e$, as shown in Fig. 6. Each fuzzy inference
system can be designed in isolation when other fuzzy inference processes are not in operation, so that the interaction from other loops is minimal. Expert experience and knowledge are essential for constructing these operating rules. The adaptive factor $\eta$ follows (3).

A separate fuzzy decoupling system is designed to suppress the coupling between these fuzzy inference systems. Similar to the autosearch system, the decoupling rules can be developed as:

- Coarse adjustment: If $\Delta\gamma^c$ or $\Delta a^c$ are relatively large, then stop the inference $F^c$. 
- Fine adjustment: If $\Delta\gamma^c$ and $\Delta a^c$ are relatively small, then start the inference $F^c$ to adjust $Q^c$.
- Regular adjustment: Start inference $F^c$, in other circumstances.

If the coupling is more complex, a more advanced strategy should be constructed in a similar manner.

**Fine Compensation**

When extensive data are available, the data-based model can be used to further improve performance, if necessary. The radial basis function (RBF) network is a good candidate because of its fast convergence and good approximation of multivariable functions [28]. The RBF network is a two-layer neural network with a nonlinear “activation” function on the hidden layer and a linear combination carried out on the output layer. The mapping $R^n \rightarrow R^r$ of the RBF is described as follows:

$$Y(X) = WZ(X) = \sum_{i=1}^{m} w_i \psi_i \left( \|X - C_i\| \right) \left( j = 1, \ldots, r \right)$$

(14)

where $X \in R^n$ is the network input vector, $\psi_i$ is the $i^{th}$ radial basis function (1 ≤ $i$ ≤ $m$). $\| \|$ denotes Euclidean distance, $C_i \in R^n$ is known as the RBF center, $w_i (1 \leq j \leq r, 1 \leq i \leq m)$ are the connection weights, $r$ is the number of output units, and $m$ is the number of basis functions. If $\psi_i$ is the Gaussian function,

$$\psi(x) = \exp \left( -\frac{x^2}{2\sigma^2} \right)$$

(15)

then the output of the hidden layer is computed as

$$z = \psi \left( \|X_p - C_i\| \right) = \exp \left[ -\frac{1}{2} \sum_{i=1}^{m} \frac{1}{\sigma_i} (X_{p_i} - C_i) \right]$$

(16)

where $z_i$ is the output of basis function $i$ at the $p^{th}$ sampling data point, $X_{p_i}, C_i$ are components of the vector $C_i$, and $\sigma_i$ is the “width.”

### RBF Learning Model (Data Based)

The simplified structure of the RBF controller is shown in Fig. 7. The linear model for $N$ in (9) is still maintained, whereas other models for $\Delta a$, $\Delta\gamma$, and $\Delta Q_e$ in (7), (8), and (10) are replaced by RBF model NN1, with the desired process variables $\Delta T_{dc}$ and $\Delta T_{uc}$ (normally, both of them are zero) as inputs. Since the cooling process is unknown, there is the Jacobian problem [29] for direct learning of NN1. A second network, NN2, is thus required to produce the estimated variables $\Delta a$, $\Delta\gamma$, and $\Delta Q_e$. The structure in Fig. 7 refers to the generalized inverse learning scheme [30]; the structure of the inverse model controller NN1 is the same as the identifier NN2. The process outputs are used as inputs to the network NN2, and the NN2 outputs are compared with the outputs of NN1 to generate the error for training network NN2. This structure will tend to force network NN2 to be the inverse of the plant.

The learning process of the RBF model attempts to identify the inverse model of the nonlinear laminar cooling process. The input and output matrices of network NN2 are

$$X = [\Delta d \quad \Delta T_m \quad \Delta T_{cm} \quad \Delta T_{dc} \quad \Delta T_{uc}]^T$$

and

$$Y = [\Delta a \quad \Delta\gamma \quad \Delta Q_e].$$

The traditional weighted RLS method [27] is employed as a learning algorithm for the RBF model. Unlike a back-propagation algorithm, only output weights are tuned when learn-
Learning of RBF Centers: The centers of the RBF network for the laminar cooling system can be chosen from the real input and output of the control system. The $k$-means clustering algorithm \cite{31} can be used to determine the centers of RBF network $C_i$.

1) Initialization: Choose initial values of RBF centers $C_i(1 \leq i \leq m)$ and the learning rate $\varepsilon(0 < \varepsilon < 1)$;

2) Similarity matching: Find the best-matching center $k$ for input vector $X(t)$ by using a minimum-distance Euclidean criterion:

$$a_i(t) = \|X(t) - C_i(t-1)\|, \quad 1 \leq i \leq m$$

$$k = \arg \min_{1 \leq i \leq m} \{a_i(t)\}$$

(17)

3) Updating: Adjust the RBF centers with the update rule

$$C_i(t) = C_i(t-1), \quad 1 \leq i \leq n, i \neq k$$

$$C_k(t) = C_k(t-1) + \varepsilon \left[ X(t) - C_k(t-1) \right]$$

(18)

Steps 2 and 3 are repeated until the proper centers $C_i$ are obtained. The $k$-means clustering is based on a linear learning rule and thus ensures rapid convergence.

Industrial Experiments

In the DCS-controlled cooling process, the supervisory control is classified into two phases: phase 1, searching for the proper set-points according to initial boundary conditions; and phase 2, maintaining consistent performance throughout the whole batch operation. If there are no corresponding set-points in the database for the new batch, the new operation mode needs to be selected. The traditional mode of supervision is manual.

To show the superiority of the hybrid intelligent control strategy, real experiments in a DCS-controlled laminar cooling process were conducted. The common boundary conditions of the slab for low-carbon steel (0.23%C) are shown in Table 3, where the first six variables, $T_{cin}$, $T_{cm}$, and $\Delta T_{du}$, are considered as disturbances to the process. The last three variables, $T_f$, $\Delta T_{io}$, and $\Delta T_{ao}$, are control requirements.

**Phase 1: Initial Search for Proper Set-Points**

When there are no corresponding set-points in the database, the fuzzy autosearch mechanism in Fig. 4 is used to explore the proper set-points for the required operating mode. The inference mechanism is traditional Mamdani type with triangular membership functions. Rules for each inference engine $F_{\gamma}$, $F_v$, $F_N$, and $F_Qe$ are extracted from the fundamental mechanism of the process and knowledge of operators and are shown in Appendix 2. The coarse compensation in (5) and (6) is used to suppress the boundary variations.

The ideal final temperature $T_{c0}$ is 600 °C, as given in Table 4. Once the actual final temperature $T_{cm}$ of the slab reaches the ideal, the proper set-point is determined. The experimental results in Fig. 8 show that the fuzzy autosearch mechanism is able to find the proper operating point after only eight steel slabs. The new operating point obtained is also given in Table 4, where the first two items are index variables, the third is the boundary variable, and the last five items are nominal set-points for control variables.

**Phase 2.1: Operation with Coarse Compensation**

After obtaining proper operating points, the system is working in phase 2.1, where the autosearch system at the higher level is switched off. This is equivalent to the absence of a human operator. The cooling process is running only in the DCS with the coarse compensation model. The final cooling temperature for most of the slabs can be maintained within the range of 600±20 °C, as shown in Fig. 9. The strictly knowledge-based compensation maintains reasonable performance for a new process when no data are available.
Phase 2.2: Operation with General Compensation

After collecting some data, the general compensation model in Figs. 5 and 6 can be used to improve performance. The rules for $F_r$, $F_a$, and $F_0$, are extracted from extensive experiments and operator experience. Since the rules are very similar to those used for the autosearch mechanism, they can share the same rule base (shown in Appendix 2) to make the implementation simpler. The only adjustment required is input and output normalization before using the rule base. The application results demonstrated in Fig. 10 show that the final cooling temperature $T_{cm}$ can be maintained within the range of 600±15 °C. The hybrid knowledge/data-based approach can achieve better performance than the strictly knowledge-based approach when a certain amount of experimental data is available.

Phase 2.3: Operation with Fine Compensation

After accumulating enough data, the strictly data-based approach can further improve performance, if necessary. Practically, the process performance accuracy is required to reach some desired level before this method can work. In an industrial environment, it may take a few months before all these conditions have been met. The experimental results in Fig. 11 show that the final cooling temperature can be controlled in the range of 600±10 °C.

Comparison Between Intelligent and Human Supervision

If there are no corresponding set-points in the database for the new batch, the classical method is to rely on a human operator to find out the new operation mode. The human operator will do a number of experiments and establish a table or a simple model, based on which he or she selects the suitable set-point for the batch. The human supervision requires extensive work, and the operating performance obtained relies on the experience of the operator. Compared to the fuzzy autosearch mechanism, the most experienced human operator may take a similar number of trials to find a suitable set-point, whereas the inexperienced operator will take many more trials. The human operator can only control the initial set-point; once the set-point is selected in this classical method, there is no other way to adjust it for the rest of the slabs in the batch. The proposed method has online compensation for the set-point that has been chosen. Therefore, the variation of the final temperature $\Delta T_{cm}$ that human supervision can maintain is around 25 °C, which is not better than the coarse compensation and worse than the general and the fine compensation methods proposed.

Conclusions

The key problem with the DCS-based laminar cooling process is that the final performance is not under closed-loop control. Instead, a human operator or similar supervision is required to adjust the set-point from time to time. Such processes exist everywhere in batch-type production, especially in the metallurgical and chemical industries, such as furnace, fermentation, and gas production plants where DCSs are being used. Even in the high-tech industry, many processes such as epoxy dispensing and snap curing in semiconductor packaging have similar features, with the final performance out of direct control. Supervision may be required to maintain consistent performance. All these examples reveal a common problem in industrial processes:
no matter how well the low-level controller performs, high-level supervision is often required to maintain good overall performance that is usually beyond the capability of direct machine control.

A hybrid control methodology combining conventional and intelligent techniques has been introduced to replace human supervision for a DCS-controlled laminar cooling process. The industrial experiments show the improved performance of the proposed hybrid control model and confirm its validity in a real manufacturing environment. The results can be extended to a wide range of processes with similar features, because the methodology presented shows superiority in several aspects due to the use of a hybrid structure:

- **High performance:** The nominal performance is maintained by the primary controller. Increased performance can always be achieved by the auxiliary compensation system. The control accuracy and the self-adaptive capability are improved gradually when the compensation model is changing from coarse to fine.

- **Reliability:** Failure of the intelligent mechanism in the auxiliary compensation model will not affect the nominal performance maintained by the primary controller.

- **Simplicity:** Construction of the auxiliary model is based on an approach similar to the piecewise strategy for nonlinear systems. This hybrid feature makes it relatively simple to develop the auxiliary compensation model by using various advanced control techniques.

- **Realization:** The suggested methodology can be easily constructed for complex industrial processes due to its hybrid strategy for different parts of the control system. The operating points of the primary controller can be obtained either from previous data, if the process has run for a long time, or from the autosearch mechanism, if the process is new. The auxiliary compensation model can be developed with simple or advanced theories, depending on the accuracy required.

Acknowledgment

The project described is supported in part under a Hong Kong government CERG grant (9040507) and a China NSF grant (69825106). The authors would like to thank the anonymous reviewers for their comments and the editors for their efforts in editing and polishing the article.

References


Appendix 1
Derivation of Approximate Linear Model for General Compensation

From (1), we know that
\[ V_T = \frac{\Delta T_v}{N} \times \frac{\Delta N_1}{N_0} \times \Delta l \]  
(A-1)

where \( \Delta l \) (\( \Delta L = N \times \Delta l \)) is the distance between two neighboring cooling units.

Because acceleration \( a \) is very small in practice (5–15 mm/s\(^2\)), its influence on speed \( \nu \) can be neglected with \( \nu = \nu_0 \). The changes in operating units \( \Delta N_1 \) due to the slab entry temperature perturbation \( \Delta T_m \) should be calculated first. From (A-1),
\[ \frac{\Delta T}{N} = \frac{\Delta T_0}{N_0} \text{ and } \Delta N_1 = N - N_0 = \frac{\Delta T_0}{\Delta T_0} \times N_0. \]  
(A-2)

Since \( \Delta T = T_m - T_w \) and \( \Delta T_0 = T_{w0} - T_{c0} \), then
\[ \Delta N_1 = \frac{T_{w0} - T_{w0}}{\Delta T_0} \times N_0 = \frac{\Delta T_{w0}}{\Delta T_0} \times N_0. \]  
(A-3)

The changes in the operating units \( \Delta N_2 \) can be obtained in terms of the slab thickness variation \( \Delta d \) [14]:
\[ \Delta N_2 = \frac{\Delta d}{d_0} N_0. \]  
(A-4)

As described previously, variables \( N \) and \( Q_e \) are strongly coupled. Variable \( N \) is a discrete integer and has rough influence on \( T_{cm} \). Variable \( Q_e \) is a continuous variable and can provide fine adjustment to performance. Therefore, \( N \) should be tuned first to counter the influence from disturbances \( \Delta T_m \) and \( \Delta d \), and the remaining uncertainties should be handled with the continuous variable \( Q_e \). If the speed \( \nu \) is constant, the changes in cooling units \( \Delta N \) due to the disturbances \( \Delta T_m \) and \( \Delta d \) can be expressed approximately as
\[ \Delta N = \text{INT}(\Delta N_1 + \Delta N_2 + 0.5) = \text{INT} \left( \frac{\Delta T_m}{T_{w0} - T_{c0}} N_0 + \frac{\Delta d}{d_0} N_0 + 0.5 \right) \]  
(A-5)

and the linear model of water flow \( \Delta Q_e \) can be formulated as follows:
\[ \Delta Q_e = \beta_{31} \Delta N_1 + \beta_{32} \Delta N_2 + \beta_{33} \left[ \frac{N_0 + \Delta N_1 + \Delta N_2}{N_0 + \Delta N} - 1 \right] Q_{e0} \]
\[ + \beta_{34} \Delta \gamma + \beta_{35} \Delta a. \]

That is,
\[ \Delta Q_e = \beta_{31} \Delta N_1 + \beta_{32} \Delta N_2 + \beta_{33} \frac{\Delta N_1 + \Delta N_2}{N_0 + \Delta N} Q_{e0} \]
\[ + \beta_{34} \Delta \gamma + \beta_{35} \Delta a. \]  
(A-6)

In (A-6), the first item compensates for the inaccuracy of model \( \Delta N_1 \), and the second item compensates for model \( \Delta N_2 \). The third item removes the integrating influence of \( \Delta N \), and the last two items represent the coupling from

| Rule base for \( F_r, F_a, \) and \( F_N \) |
|---|---|---|---|---|---|---|---|
| \( d \) | \( \Delta \gamma / \Delta a / \Delta N \) | HZ | MS | HMS | HM | HMB | HB | HL |
| NL | NL | NL | NL | NL | NL | NL | NL | NL |
| NB | NL | NL | NL | NL | NB | NB | NB | NB |
| NM | NL | NL | NL | NB | NB | NMB | NMB | NMB |
| NS | NMB | NMB | NMB | NMB | NM | NMS | NMS | NMS |
| NZ | NMS | NMS | NMS | NS | NS | NMB | NMB | NMB |
| PZ | PS | PS | PS | PS | PZ | PZ | PZ | PZ |
| PS | PM | PM | PMS | PMS | PS | PS | PS | PS |
| PMS | PMB | PMB | PMS | PMS | PS | PS | PS | PS |
| PMB | PB | PB | PM | PM | PMS | PMS | PMS | PMS |
| PB | PL | PL | PMB | PMB | PMB | PM | PM | PM |
| PL | PL | PL | PB | PB | PB | PB | PB | PB |

June 2001
IEEE Control Systems Magazine
47
flow rate $\gamma$ and acceleration $a$. Substituting (7) and (8) into (A-6) yields
\[
\Delta Q_e = \beta_{34} \Delta N_1 + \beta_{32} \Delta N_2 + \beta_{33} \frac{\Delta N_1 + \Delta N_2 - \Delta N}{N_0 + \Delta N} Q_{e0} \\
+ \beta_{34} \Delta T_m + \beta_{35} \Delta d
\]
(A-7)

where $\beta_{34} = \beta'_{34} + \beta'_{35}$, $\beta_{35} = \beta'_{34} + \beta'_{35}$.

**Appendix 2**

**Rules for Inference Engines $F_\gamma, F_a, F_N$, and $F_Q_e$**

These operating rules are extracted from practical experiment and operator experience. Rules for $F_\gamma, F_a$, and $F_N$ are similar, as summarized in Table 5. In practice, $F_\gamma, F_a$, and $F_Q_e$ can be shared by both the fuzzy systems in the primary controller and the general compensation model in the auxiliary compensator. If they are used for the general compensation model, the output of the rule base $F_\gamma, F_a$, and $F_Q_e$ will be $\Delta \gamma', \Delta \alpha'$, and $\Delta Q_{e'}$ instead of $\Delta \gamma, \Delta \alpha$, and $\Delta Q_e$, as given in Tables 5 and 6.

**Han-Xiong Li** received the B.E. degree from the National University of Defense Technology, China, in 1982, the M.E. degree in electrical engineering from Delft University of Technology, The Netherlands, in 1991, and the Ph.D. degree in electrical engineering from the University of Auckland, New Zealand, in 1996. He has been on the faculty of the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, since 1997. His research interests include intelligent control and automation, integrated control for industrial processes, and Internet-based intelligent maintenance. He was a Research Engineer from 1993 to 1997 at the University of Auckland and a Research Associate at the Delft University of Technology in 1992. He was also a Senior Process Engineer in 1997 for ASM, Hong Kong, a Project Manager from 1987-1988 at CITIC, Beijing, China, and a Software Engineer from 1982-1987 at Peking Analysis Center, Beijing, China.

**Shouping Guan** is with the Research Center of Automation, Northeastern University, China. He received the B.S.E. and the M.S.E. degrees in automatic control from Harbin Shipbuilding Engineering Institute in 1989 and 1992, respectively, and the Ph.D. degree in automatic control from Northeastern University in 1995. He was with the Engineering Department in the Research Center of Automation, the National Engineering and Technology Research Center for Metallurgical Industry Automation. His research interests are in process optimal control, intelligent control, monitoring and diagnosis in metals, and other batch-like industrial control applications.