1 Introduction

Identification of transfer function models of a system is required for an improved tuning of controllers. Several methods have been reported in the literature for identification of transfer function models with two, three and four parameters (pure delay system, first order plus time delay (FOPTD), second order plus time delay (SOPTD), etc.) using relay feedback approach. In this section, the basics of conventional relay feedback method and modifications in the original autotuning method are reviewed for single-input single-output systems. Excellent reviews on relay tuning methods are given by Yu (1999, 2006), Hang et al. (2002) and Wang et al. (2003). Methods of designing PI/PID controllers based on the transfer function models are also briefly reviewed.

1.1 Relay Feedback Method

Åström and Hägglund (1984) suggested the use of an ideal (on–off) relay (Fig. 1.1) to generate a sustained oscillation in the closed loop. For positive gain process, on–off relay is defined by $u = u_{max}$ if $e \geq 0$, and $u = u_{min}$, if $e < 0$. For negative gain processes, on–off relay is defined by $u = u_{min}$, if $e \geq 0$, and $u = u_{max}$ if $e < 0$. Amplitude (a) and period of oscillation (p_u) are noted from the sustained oscillation. This is a closed loop method for identification of transfer function models. The method is based on the observation that when an open loop output lags the input by π radians, the closed loop system may oscillate (Fig. 1.2) with a period P_u. The ultimate gain (K_u) and ultimate frequency (ω_u) can be calculated from the oscillatory response (Aström and Hägglund, 1984) from

$$k_u = 4h / (\pi a) \tag{1.1}$$

$$\omega_u = 2\pi / P_u \tag{1.2}$$

where 'h' is magnitude of the relay. Advantages of this method are: it is simple, time efficient and a closed loop method.

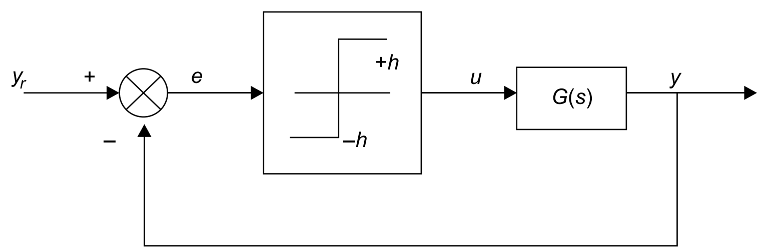

Fig. 1.1 *Block diagram for symmetric relay feedback system*

Let input to the relay be given by a sin(ωt). Then output of the relay will be a square wave, as shown in Fig. 1.2. We can approximate the square wave by a Fourier series. In general, any function can be expressed as

$$u(t) = A_0 + \sum_{n=1}^{\infty} B_n \sin(n\omega t) + \sum_{n=1}^{\infty} A_n \cos(n\omega t) \tag{1.3}$$

$$B_n = (1/\pi)\int_0^{2\pi} u(t)\sin(n\omega t)dt; \quad A_n = (1/\pi)\int_0^{2\pi} u(t)\cos(n\omega t)dt \tag{1.4}$$

Since the symmetric square wave is an odd function, $A_n = 0$ ($n = 0,1,2,\ldots$), $B_n = 0$ ($n = 0,2,4,\ldots$). Eq.(1.4) becomes

$$B_n = (2/\pi)\int_0^{\pi} h \sin(\omega t)\,dt \tag{1.5}$$

$$= 4\,h\,/\,(n\pi) \tag{1.6}$$

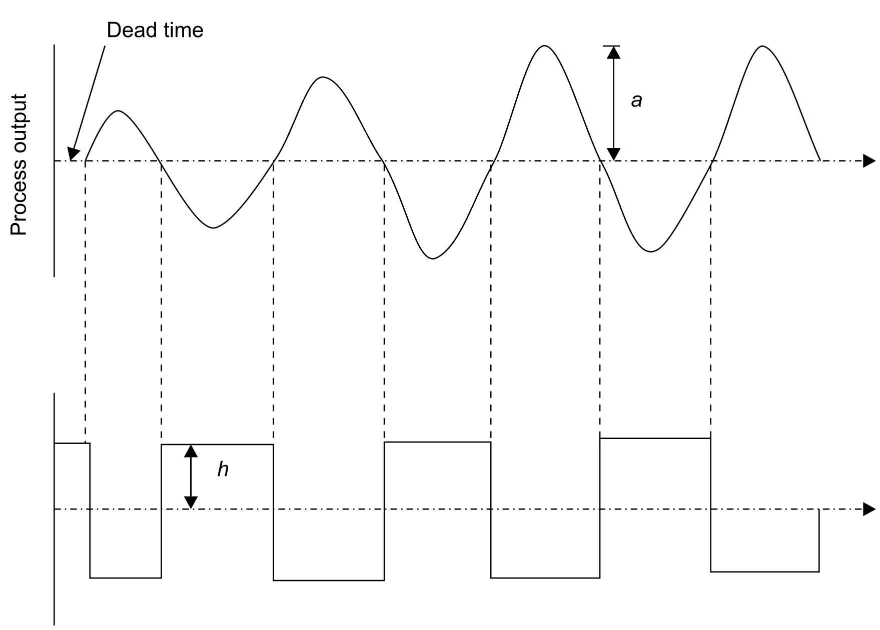

Fig. 1.2 *Response of a relay feedback system*

Hence, $u(t)$ from Eq.(1.3) can be written as

$$u(t) = (4h/\pi)\ [\sin(\omega t) + (1/3)\sin(3\omega t) + (1/5)\sin(5\omega t) + \cdots] \quad (1.7)$$

In other words, response of the relay to a sinusoidal input is a periodic signal that contains a fundamental component as well as higher order harmonics. The coefficients of higher order harmonic functions are smaller than those of the fundamental one and the higher harmonics continue to decrease as the order of the harmonics increases. Assume further that the process dynamics is of low pass characteristics and that contribution from the first harmonics dominates in the output. Error signal has the following amplitude

$$a = (4h\ /\ \pi)\ |G(j\omega t)| \quad (1.8)$$

Condition for the sustained oscillation is given by phase angle criterion

$$< G(j\omega t) = -\pi \quad (1.9)$$

and

$$k_u = 1\ /\ |G(j\omega_u)| \quad (1.10)$$

$$= 4h\ /\ (\pi a) \quad (1.11)$$

where k_u can be regarded as equivalent gain of the relay for transmission of the sinusoidal signal with amplitude a. Frequency of the limit cycle is thus automatically adjusted to the frequency ω_u, at which the open loop process has a phase lag of π radians. Physically, k_u is the ultimate gain that brings the system to stability boundary under the proportional controller action. An experiment with the relay feedback will thus give the period and amplitude of loop transfer function of the process at frequency at which the process lag is π radians. Notice that an input signal whose energy content is concentrated at ω_u is generated automatically in the relay feedback experiment. In Fig. 1.1, it is assumed that the process has a positive steady-state gain and, accordingly, the relay is defined as $u = u_{\max}$ if $e \geq 0$ and otherwise if $e < 0$. If the process has a negative steady-state gain, then the relay is defined as $u = u_{\max}$ if $e \leq 0$ and $u = u_{\min}$ if $e > 0$.

1.2 Identification by Symmetric Relay Feedback Method

Luyben (1987) used the autotuning method to identify FOPTD and SOPTD models when the steady-state gain is known *a priori*. From the initial response, value of time delay (D) is found out. Using values of k_u and ω_u in the phase angle and amplitude criteria for first order plus time delay model, values of time constant (τ) and process gain (k_p) are estimated:

$$-\tan^{-1}(\tau\,\omega_u) - (D\,\omega_u) = -\pi \quad (1.12)$$

$$k_p\,k_u\,/\,(1+\tau^2\omega_u^{\,2})^{0.5} = 1 \quad (1.13)$$

The method proposed by Luyben (1987) for obtaining simple transfer function models from a single autotune test requires the steady-state gain to be known *a priori*. This limits its usefulness for identification.

Li et al. (1991) presented a modified procedure that did not require the knowledge of steady-state gain. The method used two autotune tests. First is a normal test; second is run with an additional dead time so that the phase angle is shifted by about 45° and a smaller ultimate frequency is obtained. Then the least squares method is used to determine unknown parameters: two time constants and steady-state gain. The method is demonstrated on several simulated processes and successfully tested on an experimental heat-exchanger process. Autotune testing of both continuous and sampled-data systems is considered.

Majhi and Atherton (1999) proposed an improved method for calculating the transfer function model for the processes by relay tuning. In this method, output response is aligned with input response by shifting to left in order to note the time delay. Sets of analytical expressions are derived from a single symmetric/asymmetric relay feedback test. Using the derived expressions, exact parameters of the open loop stable and unstable FOPTD and SOPTD transfer function models are obtained.

Thyagarajan and Yu (2003) conducted relay feedback tests on processes with different orders and a wide range of dead time to time constant ratios. On the basis of the shape of response from the relay feedback test, these processes are broadly classified into three major categories. Procedures are given to find parameters for the corresponding model structures and then different tuning rules are employed to find appropriate PI controller settings. The procedure is tested against linear systems with and without noise. Several simulation results are given to show the effect of including the shape factor. Improvement is achieved from the conventional relay feedback test and no additional testing is required.

1.3 Identification using Asymmetrical Relay

Huang et al. (1996a, b) proposed a new method that uses data from a relay feedback loop to derive the SOPTD transfer function models. The relay feedback test is followed by an open loop test. Through one simple run of the proposed test, all the parameters in an SOPTD model can be identified with sufficient accuracy. The resulting SOPTD models have a close step response behavior to that of the process. The proposed method also works well in presence of a constant unknown disturbance. When SOPTD processes have a large dead time, the simulation results show that the proposed method works well with slight modifications. The disadvantage of this method is that it requires two tests – one asymmetric relay test followed by an open loop test. Also, the method cannot be applied to open loop unstable processes.

Shen et al. (1996a) proposed the input-biased relay feedback test (Fig. 1.3) for system identification. In addition to the critical point (k_u and ω_u), the steady-state gain can also be found out in a single relay test. Function analysis of the biased relay feedback is described and extensions to system identification are proposed. Potential problems associated with biased relay are given and methods to avoid them are suggested. The Autotune Variation (ATV) method is proposed for the identification of parametric transfer functions. Two linear examples are used to test the effectiveness of the proposed method. Results show that biased relay gives a satisfactory performance.

Shen et al. (1996b) proposed a saturation relay feedback test (Fig. 1.4) system to improve the accuracy of the estimates of ultimate gain and ultimate frequency. The study shows that too small a slope in the saturation relay may fail to generate a limit cycle and, subsequently, lead to a failed experiment. A procedure to overcome the trade-off between the accuracy and failure is proposed. The results show that a significant improvement can be achieved. The method first uses a conventional relay feedback and, from the test, slope of the saturation relay is calculated. Then the saturation relay test is carried out.

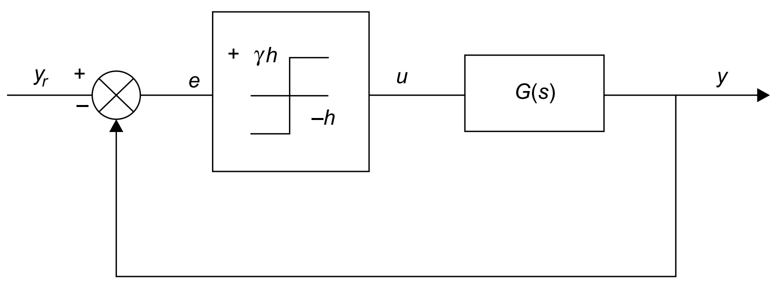

Fig. 1.3 *Block diagram for asymmetric relay feedback system*

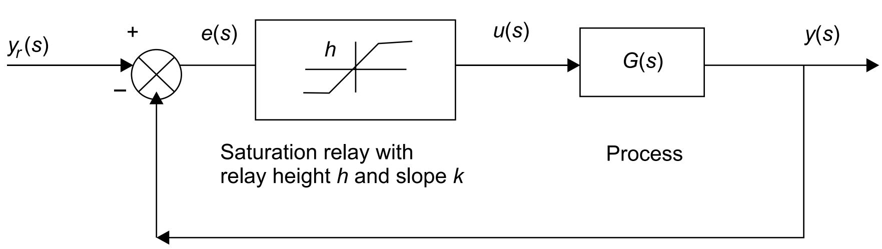

Fig. 1.4 *Block diagram for saturation relay feedback system*

Methods to improve the autotune identification methods under load disturbances are reported by Hang et al. (1993), Shen et al. (1996c), Park et al. (1997), Wang and Hwang (2003) and Leva (2005).

1.4 Identification of Unstable Processes

Kavdia and Chidambaram (1996) applied the relay tuning method for tuning controllers online for unstable processes. Response of the closed loop system with the controller settings of the identified model is compared with that of the actual system's controller parameters.

Huang and Chen (1999) proposed an autotuning procedure for PID controllers for a more general class of unstable processes that have second order dynamics. The second order dynamics are represented by models that have both stable and unstable poles together with dead time. A biased relay feedback is used to generate a constant limit cycle for identifying the model. On finishing the identification, simple tuning rules are provided to tune the parameters of PID controllers. Two unstable non-linear processes are also used to illustrate this proposed autotuning method.

Thyagarajan and Yu (2003) proposed a method to identify the approximate UFOPTD model using the shape of the response. The time constant (τ) is calculated using the equation

$$2\exp(t_{a/2}/\tau) - \exp(t_a/\tau) = 1 \tag{1.14}$$

Here t_a is the time required to reach the peak amplitude, $t_{a/2}$ is the time required to reach one half of the peak amplitude. Once τ is known, the other two model parameters (D and k_p) are obtained using the ultimate properties (k_u and ω_u) as

$$D = [-\pi + \tan^{-1}(\tau\omega_u)]/\omega_u \tag{1.15}$$

$$k_p = [1 + (\tau\omega_u)^2]^{0.5}/k_u \tag{1.16}$$

1.5 Autotuning of Cascade Control System

Cascade control system (Fig. 1.5) is a multi-loop control scheme, commonly used in chemical process industries. In cascade control systems, there are two loops: secondary (inner) loop and primary (outer) loop. The effect of disturbances entering the secondary loop is effectively controlled by the cascade control scheme. Manual tuning of cascade controllers is a tedious and time-consuming task. In view of the widespread application of relay tuning, it would be useful to automate the tuning procedure. There are two ways of tuning cascade control systems, namely, simultaneous tuning and sequential tuning procedures. Hang et al. (1994) proposed a sequential relay autotuning of cascade control loops (Fig. 1.6). The sequential tuning procedure involves two steps: The inner loop is tuned before tuning the outer loop. Using a relay for the inner loop and keeping the outer loop open, the inner loop is then tuned. The inner loop controller is designed and set and then the outer loop is kept under relay feedback in order to tune the outer loop controller.

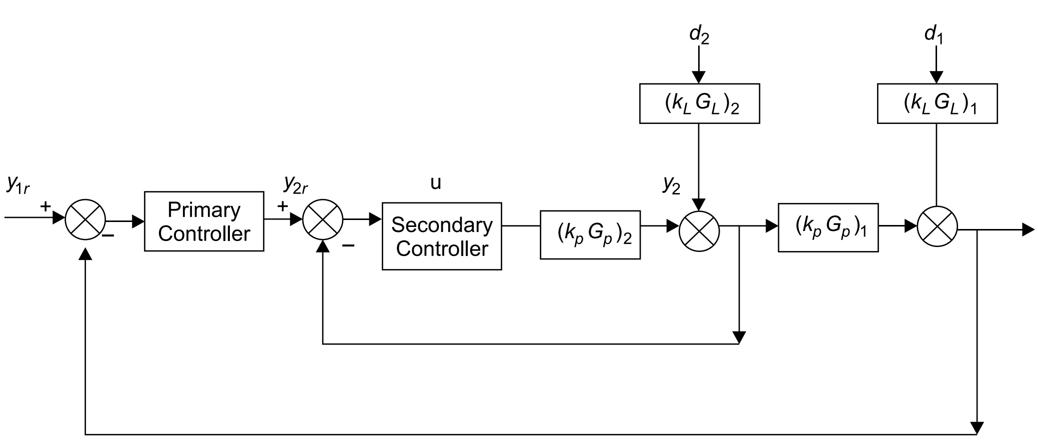

Fig. 1.5 *Series cascade control scheme*

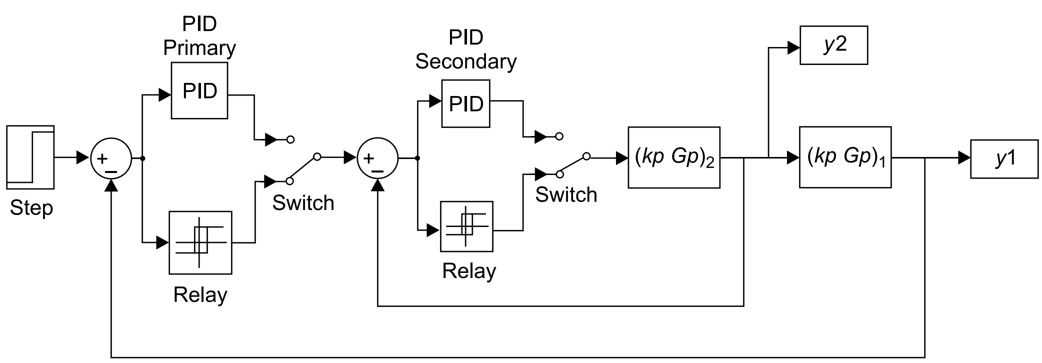

Fig. 1.6 *Relay feedback tuning of cascade control system*

Tan et al. (2000) proposed a new simultaneous online automatic tuning method for cascade control using a relay feedback approach. Departing from the traditional approach towards tuning of cascade control systems, where the secondary and primary loops are tuned in strict sequence, the proposed approach is to carry out the entire tuning process in one experiment. For ease of practical applications, the entire procedure of controller design may be automated and carried out online. A direct controller tuning approach to tune the controllers is proposed here. Robustness analysis of the new cascade control design is further carried out by drawing on existing results for SISO feedback systems. Simulation results are provided to illustrate the applicability and effectiveness of both the online autotuning approach and the new cascade control design.

Saraf et al. (2003) proposed a simultaneous autotuning method for cascade control of unstable CSTR processes. The sequential tuning method is not applicable for unstable systems. The simultaneous autotuning method is also applicable for stable systems. The method involves replacing both the controllers by the relays and carrying out the autotuning procedures to obtain limit cycles. The output of the primary auto relay loop is the input of the secondary auto relay loop. The secondary loop is forced to oscillate at the frequency of the primary loop with two relays inserted in the feedback loop because of the condition $\omega_{c,\,pri} < \omega_{c,\,sec}$. Thus the ultimate gain and ultimate period are first obtained for the primary loop. The primary controller parameters can also be obtained. Then the secondary loop executes relay feedback with the primary controller in place in order to obtain its own ultimate period, and the secondary controller parameters are obtained.

1.6 Relay Tuning of Multivariable System

Luyben (1987) used independent relay feedback method for identification of transfer function matrix of multivariable systems. The steady-state gains of each element of transfer function matrix is obtained from the step response test. Then each loop is subjected to relay feedback, keeping other loops open. From the limit cycle data, the ultimate gain and the ultimate frequencies are noted. The delays are also noted from the initial part of the responses. Then, from the amplitude and phase angle criteria, each element in the same row of plant transfer model is modeled as first order or second order system with dead time. Same procedure is used to get the other rows of transfer function matrix. Process gain and delay are assumed to be known. For each element of transfer function matrix, phase angle and amplitude criteria are given by

$$-\tan^{-1}(\tau\omega_c) - D\omega_c = -\pi \tag{1.17}$$

$$k_p k_{c,u} / (1 + \tau^2 \omega_c^2)^{0.5} = 1 \tag{1.18}$$

The equations for finding process gain and time constant for the first order plus time delay are given as

$$k_p = (1 + \tau^2 \omega_c^2)^{0.5} / k_{c,u} \tag{1.19}$$

$$\tau = (1/\omega_c)[\tan(\pi - D\omega_c)] \tag{1.20}$$

Since, this method is only a partially closed loop test, the system is more susceptible to disturbances and also is not suitable for unstable systems. Also the time delays are found from initial part of the response, which may not be accurate for systems with higher order dynamics.

Friman and Waller (1994) studied autotuning by relay feedback with an emphasis on PID control of TITO systems. They studied relay identification and autotuning in combination with integrator-plus-dead-time and gain-delay models for various systems (SISO and MIMO). In relay feedback for MIMO identification and control, relay is connected over one element in transfer function matrix and all elements in the same column can be identified from measured outputs. For gain plus delay model, the delay and gain are obtained as

$$D = 0.5 P_u \tag{1.21a}$$

$$k_p = (y_{\max} - y_{\min})/(2h) \tag{1.21b}$$

While, for integrator-plus-dead-time model, model parameters are obtained as

$$D = 0.25 P_u \tag{1.22a}$$

$$k_p = 2(y_{\max} - y_{\min})/(P_u h) \tag{1.22b}$$

After identifying the transfer function matrix, systematic controller design is carried out by model-based techniques. The approach is illustrated through a number of simulated SISO examples and two experimental 2 × 2 systems (a mixing tank and a distillation column). Other references in the MIMO systems are given by Semino and Scali (1996; 1998) and Shen and Yu (1994).

1.7 PI/PID Controller Design

For many of the control problems, a satisfactory performance is obtained using PI/PID controllers. Excellent reviews on the design of PID controllers are available (Ang et al., 2005;

Aström and Hägglund, 1995, 2001; Cominos and Munro, 2002; Datta et al., 2000; Johnson and Moradi, 2005; O'Dwyer, 2003; Sung and Lee, 1996; Tan et al., 1999; Visioli, 2005). There are special journal issues brought out on PID control (Aström et al., 2001; Isaksson and Hägglund, 2002). For controller design purposes, many processes are adequately represented by first order plus time delay models. The methods available for design of PID controllers can be broadly classified into:

(i) methods based on stability analysis (Aström and Hägglund, 2004; Chidambaram, 1998; Cohen and Coon, 1953; Ziegler and Nichols, 1942), (ii) methods based on gain and phase margin method (Wang et al. 1999a, 1999b; Wang and Shao, 1999; Wang and Cai, 2002), (iii) synthesis method (Chen and Seborg, 2002; Smith and Corripio, 2001), (iv) pole placement method (Clement and Chidambaram, 1997a), (v) IMC method (Abbas, 1997; Chien and Fruehauf, 1991; Lee et al., 1998a; Rivera et al., 1986), (vi) equating coefficient method (Padmasree and Chidambaram, 2005), (vii) optimization method (He et al., 2000; Pedret et al., 2002; Syrcos and Kookos, 2005; Toscano, 2005; Visioli, 2001) and (viii) other methods.

Foley et al. (2005) compared the performance of PID controller settings proposed by Rivera et al. (1986), Skogestad (2003), Wang and Shao (2000) and Chen and Seborg (2002). The correct choice of tuning method depends on the process control objective as well as the plant dead time to time constant ratio. WS method is recommended when the response to set point changes is of greatest concern. For purely regulatory applications, CS method gives the best overall results for lag-dominant processes ($D/\tau < 0.35$). IMC-improved PI (Rivera et al., 1986) method is better suited for the regulation of delay dominant systems ($D/\tau > 0.35$). When set point following and disturbance rejection are of roughly equal importance (as for slave loops in a cascade structure), IMC method proposed by Skogestad (2003) should be used if D/τ is small and Rivera et al. (1986) method for large values of D/τ.

Methods of designing PI/PID controllers for stable systems with inverse response systems are given by stability analysis (Ziegler and Nichols method, 1942), IMC method (Scali and Rachid, 1998; Skoestad, 2003; Wang et al., 2001), gain and phase margin method (Luyben, 2001), synthesis method (Padmasree and Chidambaram, 2005) and equating coefficient method (Padmasree and Chidambaram, 2005).

The design of PID controllers for unstable FOPTD model has attracted attention recently (Chidambaram, 1997; Padmasree and Chidambaram, 2005). The performance specifications that are usually obtained for stable FOPTD model may not be obtained for unstable FOPTD systems (Stein, 2003). The methods of designing PID controllers for unstable FOPTD systems are given by (i) modified Ziegler–Nichols method (DePaor and O'Malley, 1989; Venkatashankar and Chidambaram, 1994; Ho and Xu, 1998), (ii) IMC method (Rotstein and Lewin, 1991; Lee et al., 1998a; Marchetti et al., 2001), (iii) pole placement method (Clement and Chidambaram, 1997b), (iv) optimization method (Manoj and Chidambaram, 2001; Visioli, 2001), (v) synthesis method (Chandrasekhar et al., 2002), (vi) two degrees of freedom method (Jacob and Chidambaram, 1996; Huang and Chen, 1997; 1999), (vii) equating coefficient method (Padmasree et al., 2004). In many of these methods, one or two adjustable parameters are used to calculate the PID settings.

Padmasree et al. (2004) compared the performances of PID controller design proposed for FOPTD systems by Huang and Chen (1999), Rotstein and Lewin (1991), Visioli (2001) and their proposed method. The performance of Padmasree et al. (2004) is the best.

The important formulae used for controller tuning are as follows:

(1) Ziegler–Nichols closed loop tuning method (Ziegler and Nichols, 1942)

P-control: $k_c = 0.5\ k_u$ (1.23a)

PI-control: $k_c = 0.45\ k_u$; $\tau_I = p_u/1.2$ (1.23b)

PID-control: $k_c = 0.6\ k_u$; $\tau_I = p_u/2$; $\tau_D = p_u/8$ (1.23c)

where, k_u – ultimate gain; p_u – period of oscillations

(2) IMC tuning equations (Luyben, 2001)

$$\lambda = \max\ (1.7\ D, 0.2\ \tau) \tag{1.24a}$$

$$k_c\ k_p = (2\ \tau + D)/\ 2\lambda \tag{1.24b}$$

$$\tau_I = \tau + (D/2) \tag{1.24c}$$

where λ is the tuning parameter

(3) Simple IMC PID settings (Skogestad, 2003)

Series PID form: $c(s) = k_c\ [(\tau_I s + 1)\ /\ \tau_I s]\ [\tau_D s + 1]$

Process	$y(s)/u(s)$	k_c	τ_I	τ_D
First	$k_p \exp(-\theta s)/(\tau s+1)$	$\tau_I/[k_p(\tau_c+\theta)]$	min $\{\tau_I, 4(\tau_c+\theta)\}$	–
Second	$k_p \exp(-\theta s)/[(\tau_1 s+1)(\tau_2 s+1)]$	$\tau_I/[k_p(\tau_c+\theta)]$	min $\{\tau_I, 4(\tau_c+\theta)\}$	τ_2

Where τ_c- tuning parameter (for fast response and robustness $\tau_c = \theta$)

(4) Ideal form of PID controller (Skogestad, 2003)

Parallel form: $c(s) = k_c\ [1 + (1/\ \tau_I s) + (\tau_D s)]$ (1.25)

$$k_c = k_c[1 + (\tau_D\ /\tau_I)];\ \tau_I s = \tau_I\ [1 + (\tau_D\ /\tau_I)];\ \tau_D = \{\tau_D\ /[1 + (\tau_D\ /\tau_I)]\} \tag{1.26}$$

Process	k_c	τ_I	τ_D
$\tau_1 \le 8\theta$	$0.5\ (\tau_1+\tau_2)/(k_p\ \theta)$	$\tau_1+\tau_2$	$\tau_2/[1+(\tau_2/\tau_1)]$

Contd.