# SIMPLE MODEL-BASED PID AUTOTUNERS WITH RAPID RELAY IDENTIFICATION

**Alberto Leva** [1]

*Dipartimento di Elettronica e Informazione, Politecnico di Milano*
*Via Ponzio, 34/5 - 20133 Milano (Italy)*

Abstract: Based on some relationships between model- and relay-based PI/PID tuning, some very simple synthesis procedures are proposed that couple the advantages of model-based methods to the simplicity and clarity of relay experiments. These procedures are fast, as they employ a single relay test, and do not introduce large process upsets. They are also very light from the computational standpoint, thus well suited for low-end industrial regulators. Both simulation examples and a laboratory test are reported. 

Keywords: PID control, autotuners, model-based control, relay-based autotuning, industrial control.

## 1. INTRODUCTION

In the autotuning context, two are the main advantages of the model-based approach: a process model is available to forecast the tuning results, and specifications can be stipulated with reference to that model (Åström and Hägglund, 1995; Leva, 2001). This increases the interpretability of design parameters, and helps selecting a value for them (Leva and Colombo, 2001*b*).

Model-based tuning has also an inherent problem, however. No matter how the model structure is chosen (most often, a very simple one is assumed *a priori* (O'Dwyer, 2003), as tuning rules for complex models are not easy to devise (Isaksson and Graebe, 1999)) the identification method has a very relevant influence (Leva, 2001). Several comparisons between two model-based methods sharing the model structure can be reversed by just changing the identification procedure. The literature is often silent on this problem, while the author believes it to be among the major obstacles for a wide acceptance of model-based tuning in the applications.

From this point of view, relay-based tuning has a big advantage. It is maybe the only framework where there is no ambiguity on how the process data (in the simplest case, one point of the Nyquist curve) is obtained and used (Yu, 1999). In the author's opinion, this is one of the main reasons for the success, and the widespread use, of relay-based tuning.

Clearly, also relay-based tuning has shortcomings. A notorious one is the limited information conveyed by a few points of an *unknown* Nyquist curve. Much research is spent on this, see e.g. (Luyben, 2001; Panda and Yu, 2003; Thyagarajan and Yu, 2003; O'Dwyer, 2003). Another one, less frequently addressed, is that agreeing the control specifications (typically, the cutoff frequency and the phase margin) is not always easy and intuitive. A suitable cutoff frequency may be guessed from the relay experiment, but this is sometimes inadequate(Besançon-Voda and Roux-Buisson, 1997; Leva, 1993). The phase margin is often difficult to relate *a priori* to the desired

[1] Voice (39) 02 2399 3410, Fax (39) 02 2399 3412, e-mail leva@elet.polimi.it

closed-loop behaviour. It is not easy to solve these problems in a manner suitable for a commercial product, and in applications a lot of heuristics is used.

## 2. PURPOSE OF THIS WORK

This manuscript presents some basic results of a research aiming at establishing relationships between model- and relay-based methods. It will be shown that (a) the Internal Model Control (IMC) approach (Morari and Zafiriou, 1989) can be exploited in such a way to ease the use, and improve the results, of PI/PID tuning methods using one point of the Nyquist curve, and (b) relay feedback can circumvent the problem of accounting for the identification method—a very important problem for which little help is at present available in the literature (O'Dwyer, 2003). As witnessed by the examples, the resulting tuning procedures are fast, reliable, and very simple to understand, also for non specialists. In addition, the computational effort required is very small.

The contribution of this manuscript is not to propose 'yet another couple of tuning techniques', rather to show a possible and viable way to address the problem of joining model- and relay-based tuning. In fact, though the problem extends far beyond the scope of this work, the procedures devised indicate that some relationships between the two tuning frameworks can be characterised on a rigorous basis. Much more effort is needed for a general solution: research is underway, and results will be presented in future works.

## 3. THE PI CASE

Consider a PI regulator written in the form

$$R_{PI}(s) = K\left(1 + \frac{1}{sT_i}\right), \qquad (1)$$

and suppose that a FOPDT (First Order Plus Dead Time) process model

$$M(s) = \mu_M \frac{e^{-sL_M}}{1 + sT_M} \qquad (2)$$

is available. Applying the IMC tuning method (Morari and Zafiriou, 1989) means determining $R_{PI}(s)$ as

$$R_{PI}(s) = \frac{Q(s)F(s)}{1 - Q(s)F(s)M_r(s)}, \qquad (3)$$

where $Q(s)$ is the inverse of the minimum-phase part of $M(s)$, the rational model $M_r(s)$ is obtained with a (1,0) Padé approximation of the delay, and the IMC filter $F(s)$ is first-order, i.e., $Q(s) = (1 + sT_M)/\mu_M$, $F(s) = 1/(1 + s\lambda)$, $M_r(s) = \mu_M(1 - sL_M)/(1 + sT_M)$. This leads, see e.g. (Leva and Colombo, 2001*a*) to the tuning formulæ

$$K = \frac{T_M}{\mu_M(L_M + \lambda)}, \quad T_i = T_M, \qquad (4)$$

yielding the nominal open-loop transfer function

$$L_{n,PI}(s) = R_{PI}(s)M(s) = \frac{e^{-sL_M}}{s(L_M + \lambda)}. \qquad (5)$$

Therefore, the nominal cutoff frequency and phase margin are

$$\omega_{cn} = \frac{1}{L_M + \lambda}, \quad \varphi_{mn} = \frac{\pi}{2} - \frac{L_M}{L_M + \lambda}, \qquad (6)$$

while the design parameter $\lambda$ can be given the sense of 'desired closed-loop dominant time constant' (Morari and Zafiriou, 1989). Now, suppose that a relay test provides a point $P(j\omega_o) = Ae^{j\varphi}$ of the process Nyquist curve. The rules (4) reveal in this case a very interesting property. Given $\omega_o$, $A$, $\varphi$ and a required phase margin $\varphi_m$, if (4) are used with a model in the form (2) parameterised so that $M(j\omega_o) = Ae^{j\varphi}$, and with $\lambda$ selected so that $\omega_{cn} = \omega_o$ and $\varphi_{mn} = \varphi_m$, the resulting PI is the same as if the point were moved to $e^{j(\varphi_m - \pi)}$ with the standard relay-based tuning formulæ (that in the PI case have no further degrees of freedom). In fact, requiring that $\omega_{cn} = \omega_o$, $\varphi_{mn} = \varphi_m$, $|M(j\omega_o)| = A$, and $arg\,(M(j\omega_o)) = \varphi$, one obtains

$$\begin{cases} \dfrac{1}{L_M + \lambda} = \omega_o \\ \dfrac{\pi}{2} - \dfrac{L_M}{L_M + \lambda} = \varphi_m \\ \dfrac{\mu_M}{\sqrt{1 + (\omega_o T_M)^2}} = A \\ -\arctan(\omega_o T_M) - \omega_o L_M = \varphi \end{cases} \qquad (7)$$

Solving for $(\mu_M, T_M, L_M, \lambda)$ and substituting into (4) produces the tuning formulæ

$$T_i = -\frac{\tan\left(\varphi + \frac{\pi}{2} - \varphi_m\right)}{\omega_o}, \quad K = \frac{T_M}{\mu_M(L_M + \lambda)}, \qquad (8)$$

that are easily rewritten as the standard one-point PI tuning formulæ obtained by solving the complex equation $R_{PI}(j\omega_o)Ae^{j\varphi} = e^{j(\varphi_m - \pi)}$ for $K$ and $T_i$. This result, not apparent *a priori*, has two important and useful consequences. First, a single relay test provides a model to forecast the closed-loop transients, without requiring any further experiment or information, such as the process gain; since this model is exact at the cutoff, forecasts will be sensible. Second, it is possible to use traditional relay-based tuning rules, but give specifications in terms of $\lambda$ instead of $\varphi_m$. As will be shown in the examples, this is more intuitive, and easier to understand for a non specialist. It is sufficient to solve (7) for ($\mu_M$, $L_M$, $T_M$, $\varphi_m$) instead of ($\mu_M, T_M, L_M, \lambda$), which leads to

$$\begin{aligned} L_M &= \frac{1}{\omega_o} - \lambda, \\ T_M &= -\frac{\tan(\varphi + \omega_o L_M)}{\omega_o}, \\ \mu_M &= A\sqrt{1 + (\omega_o T_M)^2}, \\ \varphi_m &= \frac{\pi}{2} - \frac{L_M}{L_M + \lambda} \end{aligned} \tag{9}$$

Moreover, if a method with no specifications is required (a very frequent choice in low-end autotuners), $\lambda$ can be made proportional to $1/\omega_o$. This is done by replacing the first relationship in (9) with $L_M = (1 - 1/k_a)/\omega_o$, and results in the simple rules

$$\begin{aligned} T_i &= \frac{\tan\left(\varphi + 1 - \frac{1}{k_a}\right)}{\omega_o}, \\ K &= \frac{\omega_o T_i}{A\sqrt{1 + (\omega_o T_i)^2}}, \end{aligned} \tag{10}$$

where $k_a$ has the meaning of 'acceleration factor', and also its effect on the tuning is more intuitive than that of $\varphi_m$. This *modus operandi* is more effective than setting $\varphi_m$ to a 'standard', fixed value—a frequently adopted choice in simple industrial solutions. If (10) are used, the parameters of the model still come from the first three relationships of (9).

In the light of the considerations above, a very simple and effective PI tuning procedure is obtained by using the process Nyquist curve point with phase -90°, that is found easily with an integrator cascaded to the relay (Åström and Hägglund, 1995; Yu, 1999). If this approach is adopted, $k_a$ can reasonably assume values in the range 1–5, as the frequency of the point with phase -90° is normally around the main process dynamics. More detailed considerations could be made on this aspect, but these are omitted here for space limitations. Suffice to say that the proposed range is adequate for virtually any real-life problem that can be addressed with a simple autotuner like the one presented here.

## 4. THE PID CASE

Consider a one degree of freedom (1-d.o.f.) real PID regulator written in the ISA form (Åström and Hägglund, 1995), i.e.,

$$R_{PID}(s) = K\left(1 + \frac{1}{sT_i} + \frac{sT_d}{1 + sT_d/N}\right) \tag{11}$$

and suppose that a FOPDT process model, in the form (2), is available. The PID (11) can be synthesised with the IMC relationship (3) by taking again $Q(s)$ is the inverse of the minimum-phase part of (2), $F(s)$ as a first-order lowpass filter, and obtaining $M_r(s)$ with a (1,1) Padé approximation of the delay, i.e.,$Q(s) = (1 + sT_M)/\mu_M$, $F(s) = 1/(1 + s\lambda)$, $M_r(s) = \mu_M(1 - sL_M/2)/((1 + sT_M)(1 + sL_M/2))$. This leads to the tuning formulæ

$$\begin{aligned} T_i &= T_M + \frac{L_M^2}{2(L_M + \lambda)}, \\ K &= \frac{T_i}{\mu_M(L_M + \lambda)}, \\ N &= \frac{T_M(L_M + \lambda)}{\lambda T_i}, \\ T_d &= \frac{\lambda L_M N}{2(L_M + \lambda)}, \end{aligned} \tag{12}$$

and to the nominal open-loop transfer function

$$L_{n,PID}(s) = \frac{(1 + sL_M/2)e^{-sL_M}}{s(L_M + \lambda)\left(1 + s\frac{\lambda L_M}{2(L_M + \lambda)}\right)}. \tag{13}$$

Defining $\gamma = \lambda/L_M$ and $\sigma = \omega L_M$, the frequency response of (13) is then written as

$$L_{n,PID}(j\sigma) = \frac{(1 + j\sigma/2)e^{-j\sigma}}{j\sigma(\gamma + 1)\left(1 + j\sigma\frac{\gamma}{2(\gamma+1)}\right)}, \tag{14}$$

whose only parameter is $\gamma$; the nominal cutoff frequency $\omega_{cn}$ depends then only on $\gamma$, and it is possible to write

$$\omega_{cn} = \frac{1}{L_M} f_\omega(\gamma). \tag{15}$$

Lengthy but trivial computations, omitted for brevity, allow to express $f_\omega(\gamma)$ exactly, in the form

$$\begin{aligned} f_\omega(\gamma) &= \frac{1}{\sqrt{2}\gamma}\sqrt{-\chi(\gamma) + \sqrt{\chi^2(\gamma) + 16\gamma^2}}, \\ \chi(\gamma) &:= 4\gamma^2 + 8\gamma + 3, \end{aligned} \tag{16}$$

that is continuous and invertible for $\gamma > 0$. Given a Nyquist curve point (i.e., $\omega_o$, $A$ and $\varphi$) and a value for $\lambda$, first $L_M$ is found by solving

$$f_\omega(\lambda/L_M) = \omega_o L_M, \quad L_M > 0 \tag{17}$$

numerically—not a difficult task, as shown in figure 1, subject to the condition

$$\lim_{L_M \to 0} \frac{df_\omega(\lambda/L_M)}{dL_M} > \omega_o \tag{18}$$

reflecting in a constraint on $\lambda$ that is illustrated by the figure itself, and can be enforced on-line (warning the operator if required). Numbers are omitted in the figure, as the fact being evidenced is qualitative. Then, $L_M$ and $\mu_M$ are computed as in (9), and the PID is tuned with the formulæ (12).

In the PID case, one-point tuning has one degree of freedom left, that in the literature is used

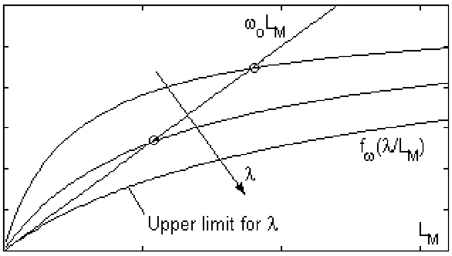


Fig. 1. An example of how $L_M$ is determined from $f_\omega(\gamma)$ and $\omega_o$.

in many different ways (Yu, 1999). Therefore, a comparison between the proposed method and one-point tuning is not very significant. The only important remark is that, also in this case, $\lambda$ is a more intuitive design parameter than a phase margin, and possibly other coefficients such as the $T_i/T_d$ ratio. Note also that the proposed method uses a *real* PID, while most one-point methods do not, and those that do normally need some further heuristics.

Here too $\lambda$ can be made proportional to $1/\omega_o$, subject to (18), $k_a$ acting as acceleration factor. Again, a simple and effective tuning procedure is obtained with the process Nyquist curve point with phase -90°, but since the PID can introduce a phase lead, also the process ultimate point can be used. The latter choice may appear preferable when load disturbance rejection is the main concern, as it inherently leads to a wider control band. If the ultimate point is used, however, experience shows that the procedure may easily result in excessive noise sensitivity. On the other hand, experience also shows that using the point at -90° and an acceleration factor $k_a$ in the proposed range allows to achieve satisfactory disturbance rejection and noise sensitivity: therefore, this is the preferred choice, and the only one used in the examples that follow.

## 5. THREE SIMULATION EXAMPLES

### 5.1 *Example 1*

This example shows PI tuning with $\lambda$ selected automatically through the acceleration factor $k_a$. The processes considered are

$$\begin{aligned} P_1(s) &= \frac{1}{(1+s)^3}, \\ P_2(s) &= \frac{1+5s}{(1+4s)^2(1+s)^2}, \\ P_3(s) &= \frac{1}{(1+100s)(1+5s)} \end{aligned} \tag{19}$$

The procedure was applied with the point at -90° and $k_a$ set to 1.2, 1.5, 2 and 4. Figure 2 shows the closed-loop step response of the process output to a load disturbance unit step. It can be appreciated that the forecasts of the transients based on the model are reasonable, as far as the main characteristics of the transients (i.e., for example, the peak value and the settling time) are considered. For a *detailed* forecast, the FOPDT structure is not adequate, but this is out of the scope of this work.

Notice that the action of $k_a$ is clear and easy to understand (the open-loop settling times of the three processes are about 6s, 12s and 300s, respectively). It is apparently simpler to relate the tuning results to $k_a$ than to the phase margin, that in this example lies approximately in the range 45°–80°.

### 5.2 *Example 2*

This example is analogous to example 1, but with the two processes

$$\begin{aligned} P_1(s) &= \frac{1}{(1+s)^2}, \\ P_2(s) &= \frac{1+s+s^2}{(1+s)^5} \end{aligned} \tag{20}$$

and $\lambda$ selected directly in the range 0.5–1.The obtained and forecast closed-loop load disturbance unit step responses are shown in figure 2. Despite the quite extreme request, the IMC-based synthesis behaves consistently, the model forecasts are reasonable, and the role of $\lambda$ remains clear enough, while the phase margins lie in the range 33°–60°.

### 5.3 *Example 3*

This example illustrates PID tuning with the point at -90°. The (moderately underdamped) process is

$$P(s) = \frac{1}{(1+s)(1+s+s^2)} \tag{21}$$

The procedure was applied with $\lambda$ set to 0.5, 1 and 1.5. Figure 4 shows the results (load disturbance step responses), and further illustrates that $\lambda$ is a 'good' tuning parameter, especially from the point of view of the potential user.

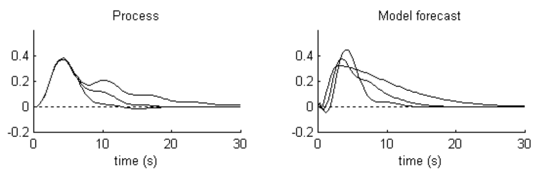


Fig. 4. Closed-loop transients in simulation example 3.

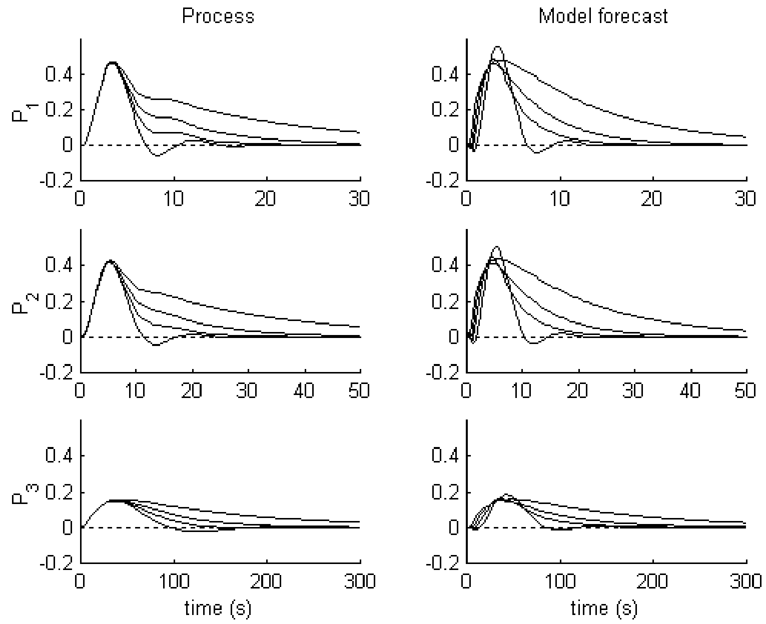


Fig. 2. Closed-loop transients in simulation example 1.

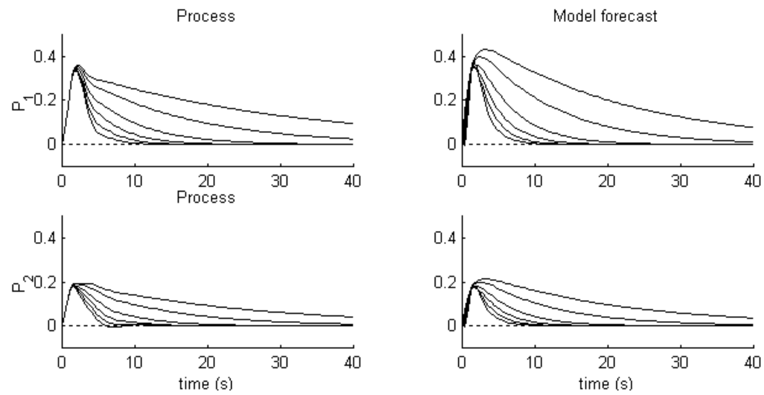


Fig. 3. Closed-loop transients in simulation example 2.

Table 1. Model and PID parameters in the laboratory test.

| $\lambda$ | $\mu_M$ | $T_M$ | $L_M$ | $K$ | $T_i$ | $T_d$ | $N$ |
|---|---|---|---|---|---|---|---|
| 5 | 0.03 | 16.7 | 7.4 | 45.7 | 18.9 | 4.0 | 1.2 |
| 10 | 0.09 | 46.7 | 2.8 | 50.2 | 47.0 | 5.6 | 0.3 |

## 6. A LABORATORY APPLICATION

The PID procedure using the point at phase -90° was applied to a laboratory apparatus where two transistors heat a metal plate, whose temperature is the controlled variable. One transistor is the control actuator, the control signal being the percentage of its maximum power, while the other provides a load disturbance. The procedure found a point with $\omega_o = 0.084$ and $A = 0.019$. The model and PID parameters obtained with $\lambda$ set to 5 and 10 are given in table 6, while the entire experiment is illustrated in figure 5.

The two FOPDT models allow also to estimate the phase margin, yielding 75.1° ($\lambda = 5$) and 78.04° ($\lambda = 10$), while a more accurate estimate, with a high order model, gives 73.58° and 79.14°, respectively. Two facts are worth noticing. First, a simple model that is precise around the cutoff allows good estimates of the tuning results. Second, should the phase margin be the specification, a small variation of it would cause a significant modification of the obtained closed-loop transients. In synthesis, then, results are satisfactory, process upset is tolerable, and - above all - the design parameter's action is clear and easy to interpret.

## 7. CONCLUSIONS

Some simple relationships between model- and relay-based PID tuning were investigated, deriving some PI/PID tuning methods aiming at coupling the advantages of model-based methods to the simplicity and clarity of relay experiments.

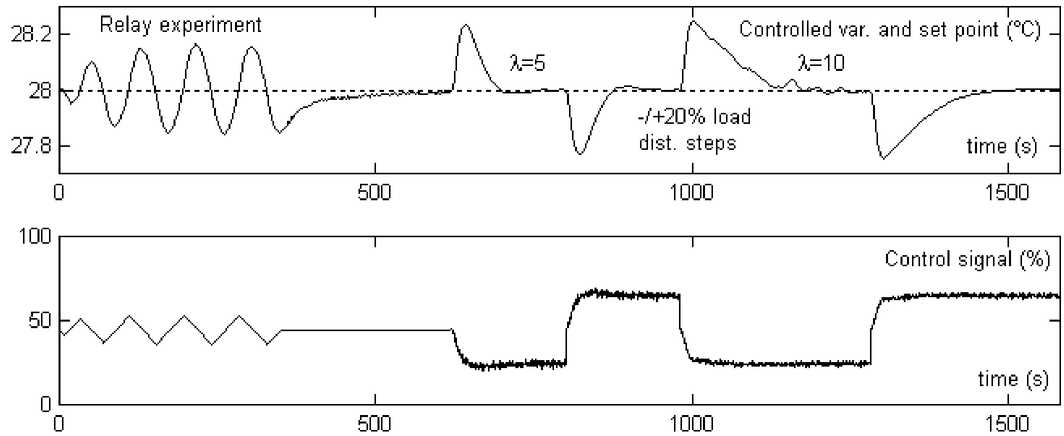

Fig. 5. Experimental transients in the laboratory test.

The rationale is that, by means of the relay experiment, the process model is made particularly precise in the band of interest for the regulator synthesis. The resulting tuning procedures are fast and reliable, requiring only a single relay test to find one point of the process Nyquist curve, introduce a very tolerable process upset, and are characterised by a single design parameter, easy to understand also for non specialists. Simulation examples and a laboratory test were reported, to demonstrate the effectiveness of the proposed approach (of which the presented procedures are just examples). Further research is being spent for a deeper analysis of the relationships between model- and relay-based PID tuning, that appear a promising framework to ease and clarify the use of existing synthesis methods, and to derive new ones. Moreover, the presented research could provide some aid for to circumvent the problem of accounting for the particular identification method used in model-based tuning, allowing to apply model-based methods in a context (the relay framework) where no ambiguity exists in the identification phase. This subject was not treated in depth due to space limitations, but it appears very interesting, and will be exploited in the next future. Finally, research is underway to exploit the generality of the idea, whose validity *per se* is not limited to PI/PID regulators.

## 8. REFERENCES


Åström, K.J. and T. Hägglund (1995). *PID controllers: theory, design and tuning - second edition.* Instrument Society of America. Research Triangle Park, NY.

Besançon-Voda, A. and H. Roux-Buisson (1997). Another version of the relay feedback experiment. *Journal of Process Control* **7**(4), 303–308.

Isaksson, A.J. and S.F. Graebe (1999). Analytical PID parameter expression for higher order systems. *Automatica* **35**, 1121–1130.

Leva, A. (1993). PID autotuning algorithm based on relay feedback. *IEE Proceedings-D* **140**(5), 328–338.

Leva, A. (2001). Model-based tuning: the very basics and some useful techniques. *Journal-A* **42**(3), 14–22.

Leva, A. and A.M. Colombo (2001*a*). IMC-based synthesis of the feedback block of ISA-PID regulators. *Proc. ECC 2001* pp. 125–247.

Leva, A. and A.M. Colombo (2001*b*). Implementation of a robust PID autotuner in a control design environment. *Trans. Inst. of Measurement and Control* **23**(1), 1–20.

Luyben, W.L. (2001). Getting more information from relay feedback tests. *Ind. Eng. Chem. Res.* **40**(20), 4391–4402.

Morari, M. and E. Zafiriou (1989). *Robust process control.* Prentice-Hall. Upper Saddle River, NJ.

O'Dwyer, A. (2003). *Handbook of PI and PID controller tuning rules.* World Scientific Publishing. Singapore.

Panda, R.C. and C.C. Yu (2003). Analytical expressions for relay feed back responses. *Journal of Process Control* **13**, 489–501.

Thyagarajan, T. and C.C. Yu (2003). Improved autotuning using the shape factor from relay feedback. *Ind. Eng. Chem. Res.* **43**, 4425–4440.

Yu, C.C. (1999). *Autotuning of PID controllers: relay feedback approach.* Springer-Verlag. London.