MODELING AND IDENTIFICATION OF NONLINEAR SYSTEMS USING SISO LEM-HAMMERSTEIN AND LEM-WIENER MODEL STRUCTURES

P. Bolognese Fernandes, D. Schlipf, J. O. Trierweiler

Abstract: This paper applies the concept of linearization around the equilibrium manifold (LEM) already presented in the literature in order to construct model structures that can be viewed as extensions of the conventional Wiener and Hammerstein models. Instead of linear time-invariant subsystems in association with static nonlinearities, these extensions exhibit variable dynamic character and can therefore model a broader class of systems than the conventional cited approaches. Moreover, the identification strategy already used with LEM systems can be applied in order to construct such models from experiments, and the techniques destined for analysis and control of Wiener and Hammerstein systems can be applied promptly. To application of these concepts to the modeling and identification is demonstrated with a numerical example, considering a heat exchange system.

Keywords: System Identification; Nonlinear Models; Linearization; Interpolation; Wiener Systems; Hammerstein Systems.

1. INTRODUCTION

In order to control satisfactorily a nonlinear plant, two main approaches exist: either the use of “inherent” nonlinear control techniques or the use of robust linear methods to guarantee stability and adequate performance in despite of the nonlinear effects. In the first approach, it is necessary that nonlinear dynamic models are available, what is very often not the case. This is mainly due to the cost of nonlinear modeling and/or identification, but also to the fact that universal and fail-free methods allowing for the identification of accurate nonlinear models are still missing.

In order to describe the nonlinear characteristics that are encountered in the practice, it is often adequate to consider a given dynamic system as the composition of a linear dynamic block followed by a static nonlinearity, the so-called Wiener system. By reversing the order of the blocks, the result is the Hammerstein model. There is a plenty of literature on specific methods for identification of either Hammerstein or Wiener models, or both. A good survey on these model structures can be found in (Pearson, 1995).

Although interesting from the practical point of view, these approaches may be too simple if the description of a nonlinear dynamics is sought. Therefore, the concept of linearization around the equilibrium manifold (LEM) can be used to include such characteristic in the model representation. The advantage of the LEM systems is that they can be
constructed in a straightforward manner and result in simple, transparent model structures.

This paper is organized as follows: Section 2 reviews the concept of LEM systems already discussed in the literature, which is the basis for the two proposed model structures. Extended Hammerstein and Wiener models, in which the dynamics are dependent of the operating point, are shown in Sections 3 and Section 4, respectively. The models are then applied in Section 5 in the modeling and identification of a nonlinear system in a numerical example. Concluding remarks can be found in Section 6.

2. LEM SYSTEMS

Consider a continuous SISO nonlinear dynamic system of the form

\[
\begin{align*}
\dot{x} &= r(x,u) \\
y &= h(x)
\end{align*}
\]  

(1)

where \( r: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^m \) is at least once continuously differentiable on \( \mathbb{R}^n \times \mathbb{R} \), and \( h: \mathbb{R}^n \rightarrow \mathbb{R} \) is at least once continuously differentiable. The output equation will be frequently omitted in the sequel for shortness. The equilibrium manifold of (1) is defined as the family of constant equilibrium points

\[
\mathcal{E} = \{(x_s,u_s,y_s) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : \quad r(x_s,u_s) = 0, \quad y_s = h(x_s,u_s)\}. \quad (2)
\]

Similarly, the family of Taylor linearizations of (1) at the set of equilibrium points determined by (2) is given by

\[
\dot{x} = \left[\frac{\partial r(x,u)}{\partial x}\right]_{x=x_s,u=u_s}(x-x_s) + \left[\frac{\partial r(x,u)}{\partial u}\right]_{x=x_s,u=u_s}(u-u_s) \quad (3)
\]

and similarly for the output equation. Under the condition that the rank of \( [\partial r(x,u)/\partial x] \) is \( n \) for all points in the set \( \mathcal{E} \) (Wang and Rugh, 1987, Fernandes 2005), the equilibrium manifold and consequently the family of linearizations of (1) will be specified by one among the \( n+1 \) variables \( (x,u) \). Therefore, if this matrix is full rank, the input fully parameterizes both families of equilibrium points and linearizations. Calling the steady-state map \( \Omega: \mathbb{R} \rightarrow \mathbb{R}^n \), such that \( r(\Omega(u),u) = 0 \) (that is, the function \( \Omega \) gives the steady-state \( x_s \) corresponding to the constant input \( u_s \)), the input-parameterized linearization around the equilibrium manifold (LEM) of (1) is defined as the system (Fernandes 2005, Fernandes and Engell, 2005).

\[
\dot{x} = A(u)(x - \Omega(u)) \quad (4)
\]

where \( A(u) \) represents the evaluation of the Jacobian matrix \( [\partial r(x,u)/\partial x] \) on \( (\Omega(u),u) \). Observe that the term arising from the second part of (3) is dropped by letting \( u = u_s \). The output equation can be linearized in an analogous fashion, considering the stationary output mapping \( \Psi: \mathbb{R} \rightarrow \mathbb{R} \). The output function \( \Omega(u) \) can be obtained on the basis of the family of parameterized linearizations by integration of

\[
\frac{d\Omega(u)}{du} = -A(u)^{-1}B(u). \quad (5)
\]

where \( A \) and \( B \) are the Jacobian matrices of \( r(x,u) \) with respect to \( x \) and \( u \), respectively, evaluated on the equilibrium manifold. The model (4) has to be interpreted as a (state-affine) nonlinear system that possesses the same family of equilibrium points (2) and the same linearization family (3) as the nonlinear system (1). Following the discussion in (Fernandes, 2005), the LEM system can constitute also a good approximation of (1) in transient regimes away from the equilibrium manifold, depending on the “degree” of nonlinearity of the original system. Obviously, other representations that are equivalent on the equilibrium manifold can be constructed on the basis of a single parameter, other than \( u \). Moreover, these representations can be easily interchanged, provided that the inverses of the corresponding elements in \( \Omega(u) \) and \( \Psi(u) \) exist.

The focus on input parameterization is due to the fact that identification experiments are carried out by exciting the plant with a designed input signal. In this sense, if one assumes that the local models can be identified by perturbing the plant around isolated equilibrium points, it is natural to use the input in order to parameterize the linearization family. Additionally, since the exact LEM system (4) involves the infinite family of linearizations and of the equilibrium points of (1), described by the matrix functions \( A(u) \) and \( \Omega(u) \), in the identification context just a finite and probably small number of the members of these families are known, but one can still use approximation or interpolation methods (for example, polynomials, splines and so on) in order to “reconstruct” these functions from the available members. Therefore, an approximation to (1) can be constructed by means of a finite number of local linear models that are considered as members of its linearization family, obtained by means of a few “local” identification experiments. In order to solve the problem of constructing a state-space representation from local models obtained from input-output experiments, these can be transformed to a linear canonical normal form prior to the constructions of the approximate functions \( A(u) \) and \( \Omega(u) \) (Fernandes and Engell, 2005). In the absence of the values of all steady-states, the last function can be obtained by integration of \( -A(u)^{-1}B(u) \) (Fernandes, 2005).

3. SISO LEM-HAMMERSTEIN MODELS

The LEM concept can be used to construct a Hammerstein-like model of (1) in which the dynamics depends on the operating point instead of
the LTI dynamics encountered in the usual Hammerstein structure (see Fig. 1).

\[
\begin{align*}
\text{Static Nonlinearity} & \quad \text{Non-linear Dynamics} \\
\theta(t) & \quad g(u(t)) & \quad w(t) & \quad G \\
y(t) & \quad \text{Output}
\end{align*}
\]

Fig. 1. LEM-Hammerstein model

A generic model with this structure can be defined in state-space form by

\[
\begin{align*}
\dot{x} &= f(x) + bw = f(x) + bq(u) \\
y &= cx
\end{align*}
\]

where \( b \) and \( c \) are vectors of proper dimensions. A possibility of constructing a model of the form of Eq. (6) on the basis of the LEM models is to separate the static nonlinear gain function from the family of transfer functions (Pearson and Pottmann, 2000), that is,

\[
G(s; \delta) = \frac{\beta_{n-1}(\delta) s^{n-1} + \ldots + \beta_1(\delta) s^1 + 1}{\alpha_n(\delta) s^n + \alpha_{n-1}(\delta) s^{n-1} + \ldots + \alpha_1(\delta) s^1 + 1}
\]

where \( \delta \) is a scalar parameterizing the set of equilibrium points/linearizations \((u_i, \text{this case}). The resulting LEM-Hammerstein system is of the form

\[
\begin{align*}
\dot{x} &= A(w)(x - \Omega(w)), w = g(u) \\
y &= cx
\end{align*}
\]

with

\[
g(u) = \int k(\delta) d\delta
\]

such that the overall family of transfer functions correspond to that of \((6). The LEM system \((8) can be constructed with realizations of the parameterized transfer function of Eq. (7) in a suitable chosen coordinate basis, as for example a canonical or normal form. Obviously, Eq. (8) depends on the new input \( w \), but an equivalent state- or output-parameterization can be easily constructed, as discussed above. These are nevertheless dynamically “worse” than the input-parameterized version (Fernandes, Engell and Trierweiler, 2004). This model can be obtained from experiments using the LEM approach as follows:

- identification of local linear models around some isolated operating points;
- transformation of the family of local models into a family of unit-gain linearizations;
- integration of \( k(u) = -C(u)A(u)^{-1}B(u) \) in order to obtain \( g(u) \);
- interpolation of \( A \) and \( B \) in some suitable canonical form and integration of in \(- A(w) B(w) \) order to generate \( \Omega(w) \).

Alternatively, since the “local” gain is the derivative of the stationary mapping with respect to \( u \) at a given operating point, \( q \) can be directly obtained by means of observations of the stationary output. This procedure can also be used iteratively, that is, values of \( y \), can be used to refine the interpolation of \( k \) and vice-versa.

4. SISO LEM-WIENER MODELS

In parallel to the Hammerstein-type structure considered above, it is also possible to define an “extended” Wiener model by replacing the linear block with an element exhibiting variable dynamics (Fig. 2).

\[
\begin{align*}
\theta(t) & \quad \text{Non-linear Dynamics} \\
G & \quad \text{Static Nonlinearity} \\
x(t) & \quad \text{Input} \\
\theta(x) & \quad \text{Output}
\end{align*}
\]

Fig. 2. LEM-Wiener model

Note that this model is not obtained by simply reversing the order of the blocks in Fig. 1, since the function \( h \) is a scalar valued function of \( n \) arguments whereas \( q \) is a \( n \)-valued function of one argument (that is, a collection of scalar functions).

This model can be defined in the state-space in the same fashion as in Eq. (6). Nevertheless, due to the nonlinear dependence of \( h \) on \( x \), the input-parameterized LEM model would exhibit a direct feedthrough characteristic, what is not desirable for simulation (Fernandes, 2005). This problem can be avoided by constructing an output-parameterized LEM system, provided that the adequate conditions hold (Wang and Rugh, 1987); in the SISO case, for example, this implies that there is no change of the sign of the stationary gain. In any case, the LEM-Wiener model is given by

\[
\begin{align*}
\dot{x}_1 &= x_2 \\
\vdots \\
\dot{x}_{n-1} &= x_n \\
\dot{x}_n &= f_n(x) + g_n(x_1)(u - \phi(x_1)) \\
y &= h(x)
\end{align*}
\]

where

\[
\begin{align*}
f_n(x) &= -\sum_{j=2}^{n} a_{j-1}(x_1)x_j \\
h(x) &= x_1 + \sum_{j=2}^{n} b_{j-1}(x_1)x_j
\end{align*}
\]

where the functions \( a_i(x_1), b_j(x_1), j, i = 0, \ldots, n-1, j \neq 0, \) correspond to the coefficients of the parameterized transfer function

\[
G(s; \delta) = \frac{b_{n-1}(\delta) s^{n-1} + \ldots + 1}{s^n + a_{n-1}(\delta) s^{n-1} + \ldots + a_0(\delta)} = g_n(\delta)
\]
where $\delta$ is a scalar parameterizing the set of equilibrium points/linearizations ($x_1$, in this case), and $a_0(x_i) = -g_i(x_i) \frac{d\Phi(x_i)}{dx_i}$. The “advantage” of this form for identification is that all involved functions are scalar and can be therefore identified by means of the variation of one single parameter. Moreover, since the output depends linearly on (10) and consequently avoids several problems.

Another advantage of the LEM-Wiener model structure is that it can be further extended by including a second-order term in the output equation, in order to improve the accuracy of the model away from the equilibrium manifold, that is,

$$h(x) = x_1 + \sum_{j=2}^{n} b_{j-1}(x_1) x_j + \Phi(x) \quad (13)$$

where $\Phi(x)$ is such that $\Phi(x_1) = 0$ and $[\partial \Phi(x)/\partial x]_{x=0} = 1$. In particular, one possibility for $\Phi(x)$ is

$$\Phi(x) = (x - \Omega(x_1))^T H (x - \Omega(x_1)) \quad (14)$$

where the $n \times n$ matrix $H$ has to be adjusted from experiments, and $\Omega(x_1) = [x_1 \ 0 \ ... \ 0]^T$. The advantage is that $H$ does not affect the dynamics of (10) and consequently avoids several problems. Moreover, since the output depends linearly on $H$, it can be adjusted by means of computationally simple methods (least-squares, for example).

5. NUMERICAL EXAMPLE: HEAT EXCHANGE SYSTEM

The model structures presented in the previous sections will be tested in the modeling and simulation of the heat exchange system (Duraiski, 2001) depicted in Fig. 3.

![Heat exchange system](image)

**Fig. 3. Heat exchange system**

This system is constituted by an insulated tank divided in three separate chambers that are allowed to transfer heat but not mass. The central chamber is in contact with both hot (h) and cold (c) chambers, but these are in contact just with the central one. The volumes of the chambers $V_h$, $V$, and $V_c$, are constant, and all chambers are well-mixed. Water is fed to and removed from each chamber separately. Under these assumptions, the system can be described by means of the following differential equations:

$$\frac{dT_h}{dt} = \frac{F_{hi} \cdot (T_h - T_i)}{V_h} - \frac{U_h \cdot A_h \cdot V_h \cdot C_p \cdot \rho}{V_h \cdot (T_h - T)} \quad (15)$$

where $T_h$, $T_c$, and $T$ are the temperatures of each chamber, $C_p$ and $\rho$ are the specific heat and specific mass of water (considered to be independent of the temperature), $F_{hi}$ and $A_h$ are respectively the overall heat exchange coefficient and heat exchange area between the corresponding chambers. A more detailed description of this system can be found in (Duraiski, 2001). In this example, the input variable is considered to be the feed flowrate of hot water, $F_{hi}$, which has constant temperature $T_{hi}$. The output is the temperature of the central chamber, $T$. The values considered for the physical parameters and other inflows can be found in the Appendix. The variation of the dynamic character is obvious from the analysis of Fig. 4 and Fig. 5.

5.1 Constructing an approximated model in LEM, LEM-Hammerstein and LEM-Wiener forms

The original LEM, LEM-Hammerstein and LEM-Wiener models described in the previous sections can be constructed analytically on the basis of the model (15). In the first case, we have a system in the form of Eq. (4) with

$$A(u) = \begin{bmatrix} -0.0033u - 0.239 & 0 & 0.239 \\ 0 & -0.0797 & 0.0797 \\ 0.239 & 0.0797 & -0.322 \end{bmatrix} \quad (16)$$

and $y = x_3$. For the LEM-Hammerstein model, it is first necessary to convert the matrices above to a normal form in order that the individual transfer functions from $w$ to $y$ in Fig. 1 have unit gain. The system is of the form:

$$\tilde{A}(u) = 10^{-2} \begin{bmatrix} 0 & 100 & 0 \\ -0.105u - 1.35 & -0.33u - 56.2 & -0.101u - 0.002 \\ 100 & 0 & -7.974 \end{bmatrix}$$

$$\tilde{B}(u) = \begin{bmatrix} w \\ 0 \\ 12.5aw \end{bmatrix} \quad (17)$$

where $u$ has to be substituted by $q^{-1}(w)$ for implementation, with

$$w = g(u) = \frac{46.2u + 376.2}{127u + 1.25} \quad (18)$$
The LEM-Wiener model (10)-(11) can be constructed similarly, giving

\[
\begin{align*}
 f_n(x) &= -0.507 \left( \frac{-0.00342 x_1 + 1.27}{(112 - 0.0304 x_1)} \right) x_2 \\
 &\quad - 0.507 \left( \frac{-0.0383 x_1 + 1.41}{(112 - 0.0304 x_1)} \right) x_3 \\
 g_n(x_1) &= 0.0238 \left( -0.0000645 x_1 \right) \\
 \phi_1 &= -7.18 \left( \frac{0.125 - 0.000418 x_1}{112 - 0.0304 x_1} \right) \\
 h(x) &= x_1 + 12.5 x_2
\end{align*}
\]  
(19)

The systems described above were simulated in Matlab with respect to the input function shown in Fig. 4; the responses are plotted in Fig. 5. The response of the linearized model at the operating point determined by \( u = 1 \) is also shown for comparison. Excepting this system, the other curves are practically indistinguishable.

The systems described above were simulated in Matlab with respect to the input function shown in Fig. 4; the responses are plotted in Fig. 5. The response of the linearized model at the operating point determined by \( u = 1 \) is also shown for comparison. Excepting this system, the other curves are practically indistinguishable.

5.2 Constructing the approximated models with identified local models

Approximated versions of the models derived in the previous section can be constructed with local models obtained either from linearizations or from identification experiments; only the last approach is exemplified here. The following procedure was adopted: three linear local models corresponding to the operating points defined by \( u_{s1} = 1 \) \( L/s \) \( (y_{s1} = 334.76 \) \( K) \), \( u_{s2} = 8 \) \( L/s \) \( (y_{s2} = 361.46 \) \( K) \), \( u_{s3} = 15 \) \( L/s \) \( (y_{s3} = 364.78 \) \( K) \) were identified by means of “local experiments”, that is, with identification signals of small amplitude around these operating points. No special methodology was employed to select the number or the location of these points; they were simply distributed over a desired range of the manipulated input. For each operating point, an identification signal \( u_{id} \) of the form depicted in Fig. 6 was designed. The switching period \( \sigma \) of the signal was determined as \( t_{63}/20 \), where \( t_{63} \) is the time needed from the step response to reach 63% of its steady-state value, what was obtained previously for each point by means of a step test with the nonlinear model (positive step of 0.2 \( L/s \) in \( u \)). The amplitude of the identification signal was fixed to 30% of \( u_{s,i} \). An input sequence of the form \( u_{s,i} - u_{id} \) was employed with validation purposes.

The response of the nonlinear model (15) was simulated in Matlab for the identification signal \( u_{id} \). In order to simulate the effect of measurement error, a white-noise, Gaussian sequence with zero mean and standard deviation of 0.01 \( K \) was added to the output. A typical plot of the noisy output measurement is given in Fig. 7. The simulated signals were sampled with a convenient sample period in order to be used with the identification algorithms.
identification data; this has the objective of reducing the effect of noise. Second, the data was filtered by means of a least-squares smoothing cubic spline (Matlab function spap2). The best parameter set of the spline function was determined iteratively in function of the results of the identification procedure achieved in the subsequent step.

The linear local models in discrete form were identified through the combined use of subspace (Matlab functions n4sid/subid) and state-space prediction error methods (Matlab function pem). The subspace methods gave the initial estimates for the prediction error method and were also used for determining the order of the state-space models. As already suggested in the literature (Fernandes, 2005), the authors found that a good local identification is generally achieved when the order of the model is clearly evidenced by the singular value test provided by the subspace routines. Moreover, a frequent indication of excessive model order and poor identification by these methods is the generation of unstable poles, complex zeros, etc. The identification procedure was as follows: first, the filtered data was used in the subspace methods; the model order was selected and the estimates were passed to the pem routine. This result was then simulated and validated against the identification and validation data. If necessary, the parameters of the smoothing spline function were changed, the data was filtered again and the local models identified once more; this procedure was repeated until a good result was found.

The local models identified in this manner were used in the construction of the model structures presented in the sections 2, 3 and 4. The LEM and LEM-Hammerstein models were constructed with local models in observability form. The last one differs from the analytical case because the linear transformation to normal form depends on the relative degree which is not a “robust” quantity to be obtained from experiments. In all cases, proper spline or rational interpolation of the necessary functions was performed (the results are omitted due to the space limitations). The responses of the three structures with identified local models for the input signal in Fig. 4 are shown in Fig 8; the agreement of the nonlinear model is quite good. The most significant difference with respect to the analytical case refers to the Wiener model, due to the identification/interpolation of the $b_i$ parameters that appear in the output function.

6. CONCLUSIONS

This paper presented new model structures based on the concept of linearization around the equilibrium manifold (LEM). These models extend the conventional Hammerstein and Wiener systems, in the sense that they allow for the inclusion of variable dynamics. These representations can be constructed on the basis of local models, which can be obtained for example by identification. A numerical example (bilinear system) showed that these structures are almost equivalent if the models are obtained analytically, but the effect of the errors in the estimated parameter can affect differently the distinct model classes.

![Fig. 8. Responses of the several systems constructed on the basis of identified local models](image)

**APPENDIX**

Parameter values used in the example:

- $\rho = 1000 \ kg/m^3$, $C_p = 4180 \ J/kg/K$, $V = V_a = V_b = 0.3 \ m^3$
- $U_kA_b = 300.000 \ J/K/s$, $U_iA_i = 100.000 \ J/K/s$
- $F_{ci} = 0 \ m^3/s$, $T_{ci} = 280 \ K$, $F_i = 0.001 \ m^3/s$, $T_i = 300 \ K$
- $T_{hi} = 370 \ K$

**AKNOWLEDGEMENTS**

The authors thank FINEP and PETROBRAS for the financial support. The first and second authors thank respectively the German Academic Exchange Service (DAAD) and the Landesstiftung Baden-Württemberg.

**REFERENCES**


