Congreso IEEE CHILECON 2015

Comparison of nonlinear identification techniques on a benchmark pH neutralization process

Juan Carlos Gómez

Laboratory for System Dynamics and Signal Processing
FCEIA, Universidad Nacional de Rosario
and CIFASIS
Rosario, Argentina

{gomez}@cifasis-conicet.gov.ar

CONICET
C I F A S I S

CHILECON 2015 - October 28-30, 2015 - Santiago, Chile

Outline

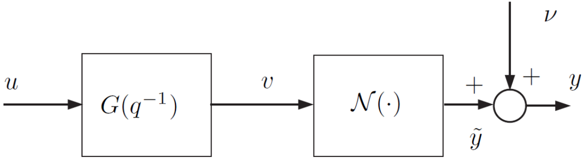
Problem Formulation

- Usually a nonlinearity test is performed before selecting the model class. Step responses around the operating point could be measured to verify if the system behaves linearly (superposition principle).
- When large deviations from the operating point are expected to occur, the predictive capability of a linear model will deteriorate, and nonlinear models will be needed to accurately represent the system.
- Even when large deviations from the operating point are taking place, it is still useful to estimate the best linear model for these operating conditions (use in MPC based on a linear model of the plant).

Problem Formulation (cont.)

Model classes

- **Best Linear Approximation (BLA) model:** optimal linear model minimizing the output prediction errors.
- **Wiener model**



- **NARX model**

$$
\begin{aligned}
y(n) &= \mathcal{F}\left[y(n-1), \cdots, y(n-n_a),\right. \\
&\qquad \left. u(n-n_k), \cdots, u(n-n_k-n_b+1)\right] + e(n)
\end{aligned}
$$

Identification Methods

- **Identification of BLA model:** based on subspace techniques (n4sid Matlab routine)
- **Wiener model ID based on orthonormal bases**
  - **Wiener model parametrization**
    - **Linear block**

$$G(q^{-1}) = \sum_{\ell=1}^{p} b_\ell \mathcal{B}_\ell(q^{-1}),$$

where $b_\ell \in \mathbb{R}$ are unknown parameters, and $\left\{\mathcal{B}_\ell(q^{-1})\right\}_{\ell=1}^{\infty}$ are rational orthonormal bases on $H_2(\mathbb{T})$.

    - **Nonlinear block**

$$\mathcal{N}^{-1}(y(n)) = \sum_{i=1}^{r} d_i f_i(y(n)),$$

where $d_i, i = 1, 2, \cdots, r$, are unknown parameters, and $f_i(\cdot), i = 1, 2, \cdots, r$ are nonlinear basis functions. Without loss of generality it is assumed that $d_1 = 1$.

Identification Methods (cont.)

- **Wiener model ID based on orthonormal bases**
  - **Wiener model ID:** By equating the values of the intermediate variable ($v(n)$) computed from the input and from the output, the following **linear regressor equation** is obtained

$$f_1(y(n)) = \phi^T(n)\theta,$$

where

$$\begin{aligned} \theta &\triangleq [d_2, d_3, \cdots, d_r, b_1, b_2, \cdots, b_p]^T, \\ \phi(n) &\triangleq [-f_2(y(n)), \cdots, -f_r(y(n)), \\ &\quad \mathcal{B}_1(q^{-1})u(n), \cdots, \mathcal{B}_p(q^{-1})u(n)]^T. \end{aligned}$$

Identification Methods (cont.)

Given an N-point data set of measured inputs and outputs $\{u(n), y(n)\}_{n=1}^{N}$, an estimate of the parameter vector θ can be computed by minimizing a quadratic criterion on the prediction errors $e(n) = f_1(y(n)) - \phi^T(n)\theta$. This is the well known **least squares estimate**, which is given by

$$\widehat{\theta} = (\Phi\Phi^T)^{-1}\Phi\mathbf{f},$$

provided the indicated inverse exists, and where

$$\begin{aligned} \Phi &\triangleq \left[\phi^T(1); \phi^T(2); \cdots; \phi^T(N)\right]^T, \\ \mathbf{f} &\triangleq \left[f_1(y(1)), f_1(y(2)), \cdots, f_1(y(N))\right]^T. \end{aligned}$$

Identification Methods (cont.)

- **Wiener model ID based on BLA**
  - **Linear block**$\rightarrow$ BLA approximation of the system
  - **Nonlinear block**

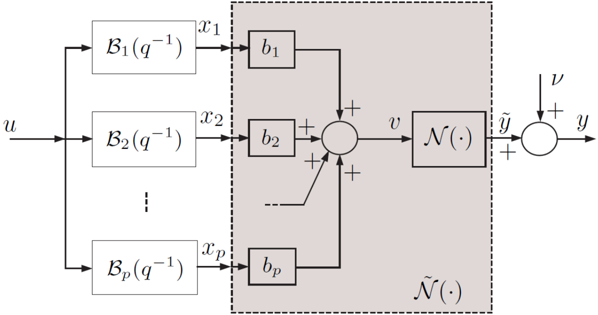
$$\mathcal{N}(v(n)) = \sum_{i=1}^{r} a_i g_i(v(n)),$$

where $a_i, i = 1, 2, \cdots, r$, are unknown parameters, and $g_i(\cdot), i = 1, 2, \cdots, r$ are nonlinear basis functions, such as polynomials, piecewise-linear functions, radial basis functions, etc..

Filtering the input with the BLA transfer function, the input to the nonlinear block is obtained. Estimates of $a_i, i = 1, 2, \cdots, r$ can be computed by **least squares fitting**, from the nonlinear block's input and output.

Identification Methods (cont.)

- **Wiener model ID based on Support Vector Regression:**



$$y(n) = \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) + \nu(n).$$

Identification Methods (cont.)

where

$$
\begin{aligned}
\mathbf{x}(n) &\triangleq \left[x_1(n), x_2(n), \cdots, x_p(n)\right]^T \in \mathbb{R}^p, \\
\mathbf{a} &\triangleq \left[a_1, a_2, \cdots, a_r\right]^T \in \mathbb{R}^r, \\
\mathbf{b} &\triangleq \left[b_1, b_2, \cdots, b_p\right]^T \in \mathbb{R}^p, \\
\mathbf{g}(\cdot) &\triangleq \left[g_1(\cdot), g_2(\cdot), \cdots, g_r(\cdot)\right]^T : \mathbb{R} \to \mathbb{R}^r, \\
\tilde{\mathbf{g}}(\mathbf{x}(n)) &\triangleq \mathbf{g}(\mathbf{b}^T\mathbf{x}(n)) : \mathbb{R}^p \to \mathbb{R}^r,
\end{aligned}
$$

Identification Methods (cont.)

- **ID problem** in the **SV regression** framework: Given a data set of measured inputs and outputs $\{u(n), y(n)\}_{n=1}^{N}$, the goal is to estimate a model of the form

$$y(n) = \mathbf{a}^T \tilde{\mathbf{g}}(\mathbf{x}(n)) + c + \nu(n),$$

where c is a *bias* term, and $\{\nu(n)\}$ is an i.i.d. random process with zero mean and finite variance. The unknowns in the model are $\mathbf{a} \in \mathbb{R}^r$, $c \in \mathbb{R}$, and the order r.

Identification Methods (cont.)

- Solution given by the **constrained optimization problem**

$$
\begin{aligned}
\min_{\mathbf{a},c,\nu} \quad & \frac{1}{2}\mathbf{a}^T\mathbf{a} + \gamma \sum_{n=1}^{N} L_\epsilon(\nu(n)) \\
\text{subject to} \quad & y(n) - \mathbf{a}^T\tilde{\mathbf{g}}(\mathbf{x}(n)) - c - \nu(n) = 0, \\
& n = 1, \cdots, N
\end{aligned}
$$

where $\gamma > 0$ is a regularization constant providing a tradeoff between model complexity (penalized by the first term in (1)) and fitting accuracy to the experimental data (penalized by the second term in (1)), and $L_\epsilon(\nu(n))$ is Vapnik's ϵ-insensitivity loss function, defined as

$$
L_\epsilon(\nu(n)) = \begin{cases} |\nu(n)| - \epsilon & \text{if } |\nu(n)| \geq \epsilon \\ 0 & \text{otherwise} \end{cases}
$$

Identification Methods (cont.)

- Introducing the **positive definite kernels**

$$K(\mathbf{x}(n), \mathbf{x}(k)) \triangleq \tilde{\mathbf{g}}^T(\mathbf{x}(n))\tilde{\mathbf{g}}(\mathbf{x}(k))$$

associated with $\tilde{\mathbf{g}}(\mathbf{x}(n))$, the **dual problem** in the Lagrange multipliers (α_n and α_n^*) can be formulated:

$$\begin{aligned}
\max_{\alpha_n, \alpha_n^*} \quad & -\frac{1}{2}\sum_{n,k=1}^{N}(\alpha_n - \alpha_n^*)(\alpha_k - \alpha_k^*)K(\mathbf{x}(n), \mathbf{x}(k)) \\
& -\epsilon\sum_{n=1}^{N}(\alpha_n + \alpha_n^*) + \sum_{n=1}^{N} y(n)(\alpha_n - \alpha_n^*) \\
\text{subject to} \quad & \sum_{n=1}^{N}(\alpha_n - \alpha_n^*) = 0 \\
& \alpha_n, \alpha_n^* \in [0, \gamma], \quad n = 1, \cdots, N
\end{aligned}$$

which is a **quadratic programming (QP) problem with box constraints**.

Identification Methods (cont.)

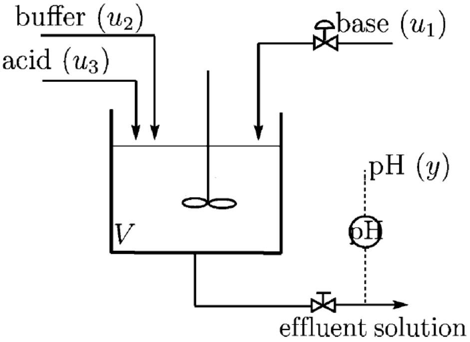
- The *dual* model representation is given by

$$y(n) = \sum_{k=1}^{N} (\alpha_k - \alpha_k^*) K(\mathbf{x}(n), \mathbf{x}(k)) + c$$

- The data points $\mathbf{x}(k)$ for which $(\alpha_k - \alpha_k^*) \neq 0$ are the so-called **support vectors**.

pH Neutralization Process

- **Input:** base flow rate, **Output:** pH of the effluent solution.



- **Acid stream:** HNO_3, **Base stream:** $NaOH$, **Buffer stream:** $NaHCO_3$

pH Neutralization Process (cont.)

- **First principles model based on reaction invariants**

$$\begin{aligned} \dot{x} &= f(x) + g(x)u_1 + p(x)u_2, \\ h(x,y) &= 0, \end{aligned}$$

where $x = [x_1, x_2]^T = [W_a, W_b]^T$, and

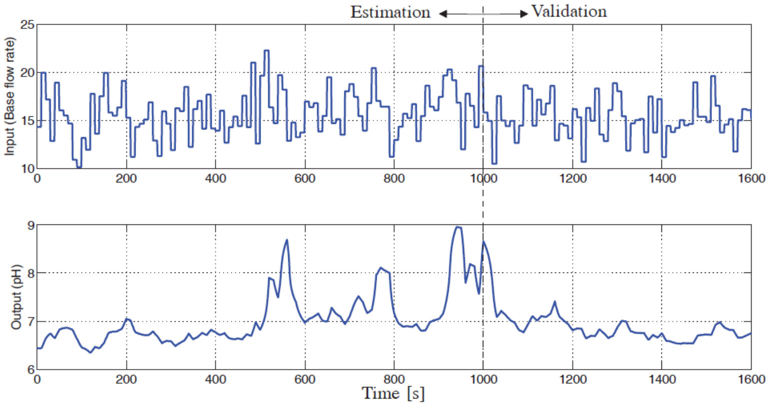
$$\begin{aligned} f(x) &= \left[\frac{u_3}{V}(W_{a3} - x_1), \frac{u_3}{V}(W_{b3} - x_2)\right]^T, \\ g(x) &= \left[\frac{1}{V}(W_{a1} - x_1), \frac{1}{V}(W_{b1} - x_2)\right]^T, \\ p(x) &= \left[\frac{1}{V}(W_{a2} - x_1), \frac{1}{V}(W_{b2} - x_2)\right]^T, \end{aligned}$$

$$h(x,y) = x_1 + 10^{y-14} - 10^{-y} + x_2 \frac{1 + 2 \times 10^{y-pK_2}}{1 + 10^{pK_1 - y} + 10^{y - pK_2}}$$

where pK_1 and pK_2 are the first and second disassociation constants of the weak acid (H_2CO_3).

Experimental Results

- **Input-Output Data:** Operating point pH = 7, output corrupted with band-limited white noise, sample period =1 sec..



Experimental Results (cont.)

- **Characteristics of the Estimated models**
  - **BLA model:** n4sid routine, 3rd. order State Space model
  - **Wiener-OB model:** 4th order linear block, OBFP with poles at $\{0.98, 0.95, 0.97, 0.60\}$, 7th. order polynomial nonlinearity.
  - **Wiener-BLA model:** BLA as linear block followed by a 7th. order polynomial nonlinearity.
  - **Wiener-SVM model:** meta-parameters of the SV-regression algorithm were set to: $\gamma = 2000$, $\epsilon = 0.001$. Gaussian Radial Basis Functions (RBF) were used as kernel functions, with a kernel bandwidth $\sigma^2 = 10$, OBFP with poles at $\{0.9849, 09849, 0.8305, 08305, 0.8305, 0.8305\}$, were employed for the linear block estimation.

Experimental Results (cont.)

- **Wiener-nlhw model:** nlhw routine of the System Identification Toolbox was used to estimate a Wiener model. A 3rd. order model was selected for the linear block, while piecewise-linear functions where selected to represent the nonlinear block.
- **NARX model:** nlarx routine of the System Identification Toolbox was used to estimate a NARX model. The following parameters were chosen for the estimation: $n_a = 3$, $n_b = 3$, $n_k = 1$. A sigmoid network was chosen as the nonlinear function $\mathcal{F}$.

Experimental Results (cont.)

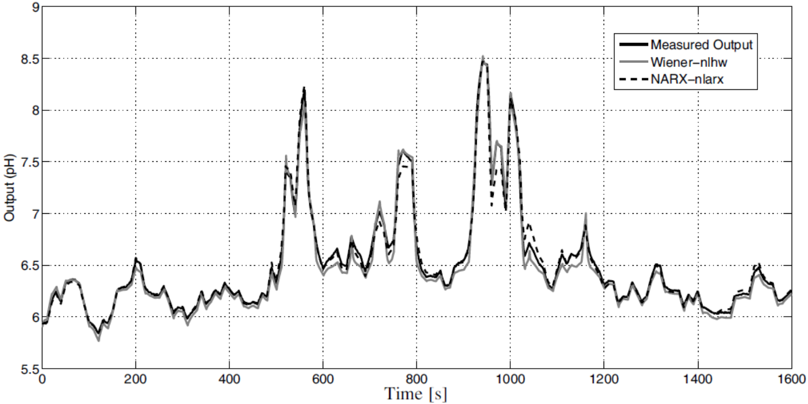
- **Prediction accuracy and number of parameters** for the different estimated models

Model	Method	Best FIT [%]	# parameters
Wiener	OB	80.3912	**12**
Wiener	BLA	83.8640	23
Wiener	nlhw	87.4905	25
Wiener	SVM	84.3179	#SV: 891
NARX	nlarx	**89.1412**	15
BLA	n4sid	60.3541	15

$$\text{Best FIT} = \left(1 - \frac{\|\mathbf{y} - \mathbf{y}_v\|}{\|\mathbf{y}_v - y_{mean}\|}\right) \times 100$$

Experimental Results (cont.)

- Validation: Measured Output (black solid line), Outputs of the estimated Wiener-nlhw (grey solid line, BestFit = 87.4905 %), and NARX-nlarx (black-dashed line, BestFit = 89.1412 %) models.



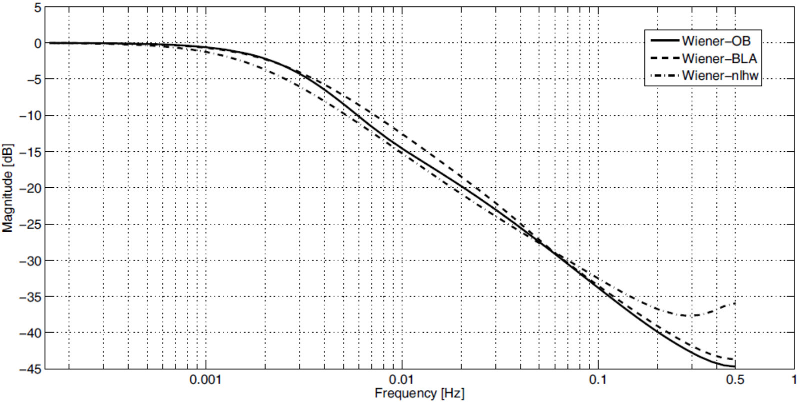
Experimental Results (cont.)

- Estimated transfer functions of the linear blocks in the Wiener models, and BLA transfer function

Model	Method	$G(z)$
Wiener	OB	$\frac{0.0066z^3-0.0163z^2+0.0130z-0.0033}{z^4-3.50z^3+4.5431z^2-2.5849z+0.5418}$
Wiener	BLA	$\frac{0.0093z^2-0.0149z+0.0063}{z^3-2.6071z^2+2.2875z-0.6793}$
Wiener	nlhw	$\frac{z-0.7834}{z^3-1.337z^2+0.08087z+0.2606}$
BLA	n4sid	$\frac{0.0093z^2-0.0149z+0.0063}{z^3-2.6071z^2+2.2875z-0.6793}$

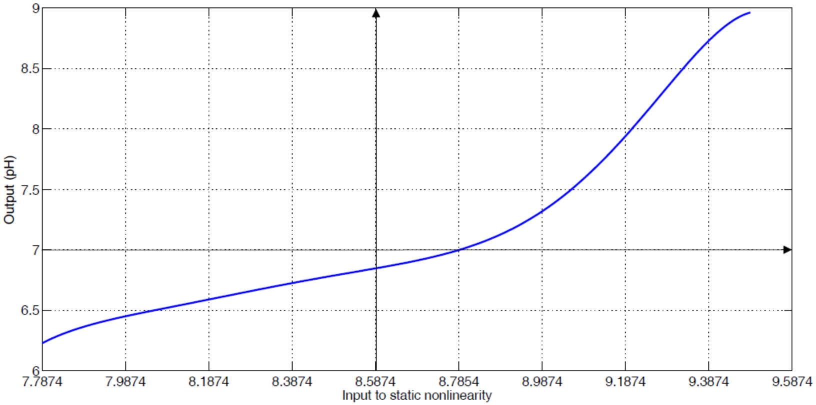
Experimental Results (cont.)

- Normalized (to unit static gain) magnitude frequency responses of the linear blocks in the estimated Wiener-OB (solid line), Wiener-BLA (dashed line), and Wiener-nlhw (dash-dotted line) models.



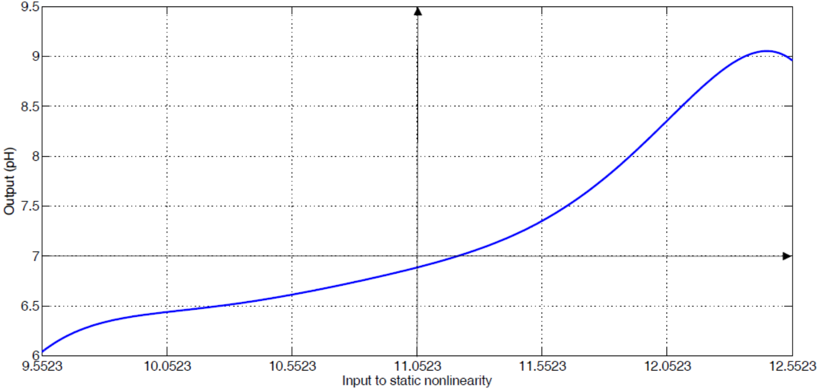
Experimental Results (cont.)

- Estimated static nonlinearity for the Wiener-OB model.



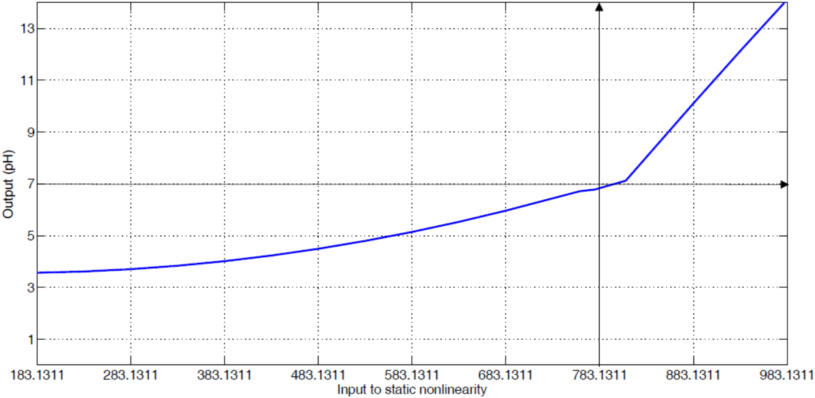
Experimental Results (cont.)

- Estimated static nonlinearity for the Wiener-BLA model.



Experimental Results (cont.)

- Estimated static nonlinearity for the Wiener-nlhw model.



Conclusions

- Several state-of-the-art nonlinear identification methods have been compared on a benchmark pH neutralization process.
- Wiener models estimated using methods based on orthonormal bases representations, as well as on methods based on support vector regression techniques were considered.
- The Wiener model with the best accuracy was the Wiener-nlhw model, but with a relatively large number of parameters.
- The Wiener model with the best tradeoff between fitting accuracy and model complexity was the Wiener-OB model.
- The overall best fitting accuracy was obtained by the estimated NARX model, with a reasonable model complexity.

Muchas Gracias !!!

Comparison of nonlinear identification techniques on a benchmark pH neutralization process

Juan Carlos Gómez

Laboratory for System Dynamics and Signal Processing
FCEIA, Universidad Nacional de Rosario
and CIFASIS
Rosario, Argentina

{gomez}@cifasis-conicet.gov.ar