PARAMETER ESTIMATION AND PREDICTION OF NONLINEAR BIOLOGICAL SYSTEMS: SOME EXAMPLES

T.G. Doeswijk * **K.J. Keesman** *

* *Systems and Control Group, Wageningen University*
P.O. Box 17, 6700 AA Wageningen, THE NETHERLANDS
Tel. +31 317 483780, e-mail: karel.keesman@wur.nl

Abstract: Rearranging and reparameterizing a discrete-time nonlinear model with polynomial quotient structure in input, output and parameters ($x_k = f(Z, p)$) leads to a model linear in its (new) parameters. As a result, the parameter estimation problem becomes a so-called errors-in-variables problem for which a total least squares approach provides a natural solution. Retrieving the predictor form after estimation leads to the modified predictor: $\hat{x} = \tilde{f}(Z, \hat{\theta})$. The objective of this paper is to evaluate the predictive quality of $\hat{x}_k = \tilde{f}(Z, \hat{\theta})$ and $\hat{x}_k = f(Z, \hat{p})$ with parameters estimated using different least squares methods. The well-known Michaelis-Menten kinetics are used as an illustration with simulated (noisy) data. Finally, an example of a storage facility containing biological products is presented with real experimental data. ©IFAC 2006

Keywords: Parameter estimation, Nonlinear identification, Calibration, Validation, Biotechnology, Total least squares

1. INTRODUCTION

In general, in a model calibration procedure, the parameters of the model are estimated such that the model predictions fit well on measured data. Several techniques are available to obtain suitable estimates. For models that are *linear* in their parameters, the ordinary and weighted least squares techniques (Norton, 1986; Ljung, 1987) are probably the most frequently used. However, when errors are not only present in the current output but also in the previous outputs and inputs, a modification of the ordinary least squares technique is required. The overdetermined set of linear equations is then given by:

$$A\theta \approx b \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$ with $m > n$ the data matrix, $\theta \in \mathbb{R}^n$ the linear parameters and $b \in \mathbb{R}^m$ the output. The total least squares (TLS) approach (Golub and Van Loan, 1980) provides a solution if errors are not only present in the output b but also in the data matrix A. The TLS technique is currently widely used in several application areas, *e.g.* Van Huffel and Lemmerling (2002) and references in Van Huffel and Vandewalle (1991), and examples are present where TLS shows its superiority over ordinary least squares.

Usually, for *nonlinear* in the parameter models, parameter estimates are found iteratively using optimization algorithms. The existence of local minima and high computational effort may impede the application of nonlinear estimation methods.

Given the nonlinear discrete-time model

$$x_k = f(Z, p) \tag{2}$$

where $Z = (x_{k-1}, \cdots, x_{k-\tau}, u_{k-1}, \cdots, u_{k-\tau})$, $k, \tau \in \mathbb{Z}^+$ and $\tau < k$ with τ the time delay, and $f(\cdot)$ a *finite polynomial quotient* in the elements of Z and $p \in \mathbb{R}^q$. The model can be rearranged and reparameterized such that a model arises that is linear in its (new) parameters θ, a polynomial quotient in p (Doeswijk and Keesman, 2005). These parameters can then be

estimated with a non-iterative least squares estimator. If noise is added to (2), then by rearranging the model, errors become part of the data matrix. Hence, in general the parameter estimation problem becomes an errors-in-variables (EIV) problem for which TLS provides a natural solution. Finally, the linear in the parameters model can be rewritten in predictor form, *i.e.*

$$\hat{x}_k = \tilde{f}(Z, \hat{\theta}) \tag{3}$$

where $\hat{\theta}$ contains the least squares estimate in the linearly reparameterized model. Note that the original parameter vector p is *not* necessarily re-estimated from $\hat{\theta}$. In what follows, the main focus is on the predictive quality of (2) and (3).

The validity of reparameterization and linear estimation was illustrated in Doeswijk and Keesman (2005) with an example of Michaelis-Menten kinetics. In the *noise free* case the linearly reparameterizing method led to the exact solution where the nonlinear least squares approach could end up in local minima. Furthermore, the linear regressive reparameterization approach was applied to a storage facility containing a biological product. Real data were used to evaluate the predictive quality of (3) with θ estimated with a truncated least squares estimator. The results were compared with the original model (2) where the parameters p were estimated by a traditional nonlinear estimation approach. The linearly reparameterized storage model, however, got an EIV structure because errors appeared in the data matrix. In a previous paper (Doeswijk and Keesman, 2005) this was neglected initially but it has been noticed that a TLS approach would be more appropriate for parameter estimation than the truncated least squares method. To our knowledge this reparameterization, analysis and application of a TLS approach to a *nonlinear in the parameter model* has never been explicitly reported.

The objective of this paper is to evaluate the predictive quality of (3) with θ estimated with OLS and TLS and compare this with the predictive quality of the *nonlinear* original model (2) with p estimated with a nonlinear least squares technique. First, the Michaelis-Menten model is used with simulated noisy data. Second, a storage model with real data is used.

2. BACKGROUND

2.1 Algebraic non-linear parameter estimation

If $f(\cdot)$ is a polynomial quotient in Z and p, reparameterizing the original model (2) leads to

$$F_0(x_k, Z) = [F_1(\cdot) \quad F_2(\cdot) \quad \dots \quad F_n(\cdot)]\,\theta \tag{4}$$

with, $F_i(\cdot) = F_i(x_k, Z)$, $i = 1, \dots, n$. The new parameters are given by $\theta_i = \varphi_i(p)$, a polynomial quotient in p. The model (4) is a linear regressor (see also (1)) and the parameters θ can be estimated by least squares.

As $F_i(\cdot)$ can be a function of x_k, even in case of an equation error structure, the data matrix may contain errors. Furthermore, the columns $F_i(\cdot)$ can become linear dependant which leads to (near) rank deficiency. Finally, nonlinearities can occur in uncertain regression variables which in turn can lead to biased estimates. Regularization and bias compensation therefore might be needed.

After estimating the parameters θ, the model can be rewritten in predictor form (3). Remark that no effort is done to estimate the original parameters p from $\hat{\theta}$ (see for details Doeswijk and Keesman (2005)).

2.2 Michaelis-Menten kinetics

A discrete-time model describing the substrate concentration in a batch bioreactor with Michaelis-Menten kinetics is given by:

$$S_k = S_{k-1} - V_{\max}\frac{S_{k-1}}{K_m + S_{k-1}} \tag{5}$$

with S the substrate concentration, $V_{\max}$ the maximum substrate conversion rate and K_m the Michaelis-Menten constant. Estimating K_m and $V_{\max}$ may lead to local minima (Doeswijk and Keesman, 2005). Rearranging (5) leads to:

$$S_{k-1}(S_k - S_{k-1}) = [-(S_k - S_{k-1}) \quad -S_{k-1}] \begin{bmatrix} K_m \\ V_{\max} \end{bmatrix} \tag{6}$$

which is linear in the parameters. Remark that in this case no reparameterization has taken place, *i.e.* $\theta = p$. As the substrate concentration is the measured (noisy) variable in this system, it can be clearly seen that the data matrix A, in this case $[-(S_k - S_{k-1}) \quad -S_{k-1}]$ for $k = 1, \dots, m$, contains measurement errors and hence, (6) has become EIV.

2.3 Storage model

A model that describes temperature dynamics in storage facilities for biological products such as fruits and vegetables, is given by (Keesman *et al.*, 2003):

$$\begin{aligned} T_{p,k} = &\left(p_1 + \frac{p_2}{p_3 + p_4 u_{k-1}} + \frac{p_5}{p_6 + p_7 u_{k-1}}\right) T_{p,k-1} \\ &+ \frac{p_8 + p_9 u_{k-1}}{p_3 + p_4 u_{k-1}} T_{e,k-1} + \frac{p_{10} + p_{11} u_{k-1}}{p_6 + p_7 u_{k-1}} X_{e,k-1} \\ &+ \left(p_{12} + \frac{p_{13}}{p_6 + p_7 u_{k-1}}\right) \end{aligned} \tag{7}$$

where $T_{p,k}$ is the measured output variable. The variables T_p denotes the temperature of the produce (^oC), T_e the external temperature (^oC) and X_e the external absolute humidity (kg/kg). Finally, the input u denotes the product of fresh inlet and ventilation and is bounded by: $0 \leq u \leq 1$. The nonlinearities in (7) are related to heat and mass transfer.

Rearranging (7) into a linear regression format leads to:

$$T_{p,k} = \begin{bmatrix} u_{k-1}T_{p,k} & u_{k-1}^2T_{p,k} & T_{p,k-1} & u_{k-1}T_{p,k-1} \\ u_{k-1}^2T_{p,k-1} & T_{e,k-1} & u_{k-1}T_{e,k-1} & u_{k-1}^2T_{e,k-1} \\ X_{e,k-1} & u_{k-1}X_{e,k-1} & u_{k-1}^2X_{e,k-1} & u_{k-1} \\ u_{k-1}^2 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \cdots \theta_{14} \end{bmatrix}^T \tag{8}$$

where $\theta_i = \varphi_i(p)$.

As can be seen from (8) the model clearly is EIV as $T_{p,k}$ is no longer only the output variable but now also appears in the data matrix. Furthermore, $T_{p,k-1}$, $T_{e,k-1}$ and $X_{e,k-1}$ are also measured uncertain variables. After estimating θ the model (8) can be rewritten in predictor form

$$\hat{T}_{p,k} = \frac{\hat{\theta}_3 T_{p,k-1}}{1 - \hat{\theta}_1 u_{k-1} - \hat{\theta}_2 u_{k-1}^2} + \\ \ldots + \frac{\hat{\theta}_{13} u_{k-1}^2}{1 - \hat{\theta}_1 u_{k-1} - \hat{\theta}_2 u_{k-1}^2} + \hat{\theta}_{14} \tag{9}$$

In what follows, the quality of the predictor (9) will be evaluated in terms of the mean square error (MSE) of the prediction errors.

3. ESTIMATION METHODS

To evaluate the performance of predictors (2) and (3) with estimates $\hat{p}$ and $\hat{\theta}$, several estimation methods are compared.

A nonlinear parameter estimation procedure for estimating the parameters of (7) is used in this paper as a reference for the alternative estimation methods. The vector p of (7) consists of several physical and design parameters. A selection of these physical and design parameters has been estimated using a nonlinear least squares procedure.

For systems linear in their parameters the ordinary least squares procedure is widely used. For ill-conditioned systems, however, the estimates are very sensitive to the data and hence, the errors therein. The predictive quality of the original system with the estimated parameters is then very poor. A regularization method to overcome these limitations is the truncated least squares method which uses the numerical rank of the data matrix to stabilize the solution (Norton, 1986, p.77).

An underlying assumption of the ordinary least squares estimator is that all errors are subjected to the output vector. However, frequently errors are not only present in the output vector but also in the data matrix. A fitting technique that compensates for errors in the data is TLS. Golub and Van Loan (1980) outlined a TLS-solution of the EIV problem which is heavily based on the singular value decomposition (SVD) of (1), *i.e.*:

$$U\Sigma V^T = [A, b] \tag{10}$$

Table 1. Nominal and estimated parameter values for different numerical ranks.

θ^0	$\hat{\theta}(\mathcal{R}=5)$	$\hat{\theta}(\mathcal{R}=4)$	$\hat{\theta}(\mathcal{R}=3)$	$\hat{\theta}(\mathcal{R}=2)$
0.9500	0.9643	0.9600	0.9527	0.9827
0.1000	2.1947	-1.7327	0.0903	0.0355
0.2000	-8.2878	-0.2826	0.1877	0.0706
0.3000	5.1155	1.1424	0.2879	0.1061

with $U \in \mathbb{R}^{m\times(n+1)}$ and $V \in \mathbb{R}^{(n+1)\times(n+1)}$ orthogonal and $\Sigma \in \mathbb{R}^{(n+1)\times(n+1)}$ diagonal, containing the singular values. Their method assumes that the errors are independent and identically distributed. Then, if the absolute size of the errors are all roughly equal (Van Huffel and Vandewalle, 1991) good results can be obtained.

Several extensions of the TLS algorithm have been proposed. In this paper the generalized TLS (GTLS) (Van Huffel and Vandewalle, 1989) algorithm is used. GTLS overcomes the limitation of independent and identically distributed errors. GTLS considers different sized and correlated errors as well as error-free variables where the covariance matrix of the errors must be known or estimated. If there are no error-free data and the covariance matrix equals the identity matrix the GTLS solution is equal to the classical TLS solution. The key problem is how to choose the covariance matrix if the error distributions are unknown.

In the storage model example, the matrix of interest for GTLS estimation, *i.e.* $[A, b]$, of the linearly reparameterized model (8) has a large condition number ($> 10^6$). The large condition number implies that some columns are highly correlated which leads to (near) rank deficiency. This, in turn, leads to an estimator that is very sensitive to the data. Hence, a regularization method was chosen. Both truncated least squares and the nongeneric GTLS use a numerical rank to overcome the problem of ill-conditioning. By choosing an appropriate numerical rank a compromise is found between the stabilization of the solution and the accuracy of the GTLS estimator.

Let us, as an illustrative example, consider the following dynamic system:

$$x_{k+1} = \theta_1 x_k + \theta_2 u_1 + \theta_3 u_2 + \theta_4 u_3 \\ y_k = x_k \tag{11}$$

where u_1, u_2 and u_3 are all constant inputs related by $u_3 = 1.5u_2 = 3u_1$. Given the set of nominal parameters, the constant input signals and the initial condition, the system is simulated for 1000 steps. All columns of the data matrix $A = [y_k, u_1, u_2, u_3] \in \mathbb{R}^{m\times n}$ and output vector $b = y_{k+1} \in \mathbb{R}^{m\times 1}$ are then corrupted with additive independent and identically distributed zero mean gaussian white noise. The matrix $[A, b]$ is now ill-conditioned due to the near linear dependency. However, the parameters can still be identified with nongeneric GTLS by evaluating the numerical rank $\mathcal{R}$ of $[A, b]$. From table 1 it can be seen that for $\mathcal{R} = 3$ the estimates approach the nominal parameters. Notice from Fig. 1 that the predictive value

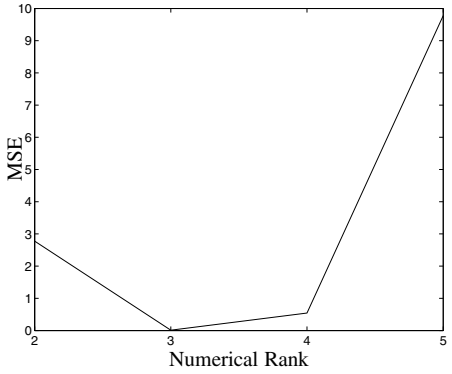

Fig. 1. MSE of simulated output with nominal parameters minus simulated output with estimated parameters for various numerical ranks

of the estimates is best in terms of MSE for numerical rank $\mathcal{R} = 3$ and the predictive performance dramatically decreases for higher (and lower) numerical rank.

The GTLS is elaborated for the bioreactor model (6) and for the storage model (8). In section 4 the predictive quality of the predictors (5) and (9) with GTLS estimates are evaluated.

4. RESULTS

4.1 Michaelis-Menten kinetics

It is known that the parameters of the Michaelis-Menten equation (5) in a batch reactor are theoretically identifiable if the initial conditions are known (Godfrey, 1983). If noise is present, however, the parameters, specifically K_m, are hard to identify (Holmberg, 1982). Furthermore, a nonlinear estimation method can suffer from local minima. In the following, a comparison between ordinary least squares and GTLS using (5) and simulation data is made. The aim of this simulation experiment is to distinguish between the predictive quality of (5) with the least squares estimates versus the GTLS estimates.

First, the substrate concentration was simulated for $k = 1, \cdots, m$, with $m = 30$ using (5) and where $K_m = 10$, $S_0 = 30$, and $V_{\max} = 2$. In addition, a Gaussian white noise sequence is generated with zero mean and unit variance. Next, the noise sequence is multiplied with a factor varying from 0.01 to 0.2 to obtain a range of standard deviations of the measurement noise. The noise corrupted substrate concentrations were then used to generate the data vectors: $-S_{k-1}$ and $\frac{S_{k-1}}{S_k - S_{k-1}}$ for $k = 1, \cdots, m$. Given these data K_m and $V_{\max}$ could be estimated. This procedure was repeated 100 times. The mean of the estimated parameters $\hat{K}_m$ and $\hat{V}_{max}$ for each standard deviation were then used for open loop prediction of the substrate concentration with (5) and $S_0 = 30$.

In order to reduce the number of data errors, the modified Michaelis-Menten equation (6) was further

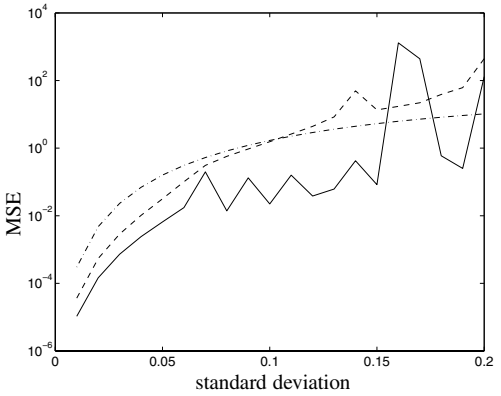

Fig. 2. MSE of predictors (5) with parameters estimated by LS of (6) ($-\cdot$) and of (12) (- - -), and by GTLS of (12) (—).

rearranged such that only one column of the data-matrix contains errors, *i.e.*

$$-S_{k-1} = K_m + \frac{S_{k-1}}{S_k - S_{k-1}} V_{\max} \qquad (12)$$

Next, for applying GTLS, the covariance matrix had to be chosen. Given the error $e_k = b_k - S_k$ the variance of the regressor of $V_{\max}$ was assumed to be much larger than the variance of the output error because the error of the denominator $e_k - e_{k-1}$ can become close to zero. The covariance matrix was chosen as:

$$\mathrm{cov}([\frac{e_{k-1}}{e_k - e_{k-1}}, -e_{k-1}]) = \begin{bmatrix} 10^6 & 0 \\ 0 & 1 \end{bmatrix}$$

Subsequently, the parameters were estimated with the least squares estimators related to (6) and (12) and with the GTLS estimator related to (12). The MSE of the prediction error ($\frac{1}{N}\Sigma_{k=1}^{N}(S_k - \hat{S}_k)^2$) is given in Fig. 2.

GTLS clearly outperforms the ordinary least squares. At higher noise levels, however, the estimates may induce singularities. At this point the ordinary least squares estimates already have a poor predictive performance.

4.2 Storage model

Let us consider again eqn (8). If the size of the errors of each column of the matrix $[A, b]$ correspond to the absolute values of the measurements then the errors are not identically distributed. The absolute humidity X_e is an order of magnitude of 3 smaller than T_e and T_p. Furthermore, the last column of the data matrix A is error free. Hence, better estimates are expected for GTLS compared to truncated least squares.

For applying GTLS a proper covariance matrix must be chosen. As the exact sizes of the errors were not known these were approximated by the variance of the columns. The correlation terms are neglected for convenience. They are relatively small and therefore will have a minor effect on the results.

Table 2. MSE of predictors with parameters estimated by nonlinear LS, truncated LS and GTLS in calibration (data set 1) and validation (data set 2) period.

	nonlinear LS	truncated LS	GTLS
calibration	0.027	0.019	0.023
validation	0.113	0.031	0.053

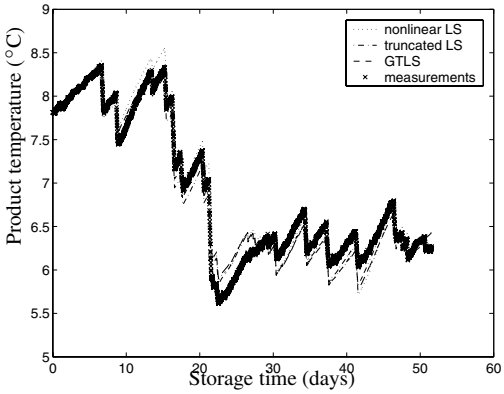

Fig. 3. Measured and predicted product temperatures in the calibration period (data set 1) with parameters obtained by different calibration techniques.

If the system is rewritten as a predictor (9) there appears to be a constraint on the estimated parameters $\hat{\theta}_1$ and $\hat{\theta}_2$. The denominator may not be equal to zero for the whole range of u. Hence, the constraint is given by:

$$1 - \hat{\theta}_1 u_{k-1} - \hat{\theta}_2 u_{k-1}^2 \neq 0, \quad 0 \leq u \leq 1 \qquad (13)$$

If the constraint is violated, the solution is rejected.

Two data-sets with measured variables, *i.e.* T_p, T_e, X_e and u, of about 50 days with a sampling interval of 15 minutes were available. The data were obtained from the same location at the same season but for a different period within the season. All parameters are assumed to be constant during the whole season. The parameters are calibrated over one data set (calibration period). The predictive quality is then obtained by using an open loop prediction over the same data set and cross-validated over the second data set (validation period).

The MSE of the predictor (7) with parameters estimated by nonlinear least squares ($\hat{p}$) and the MSE of predictor (9) with the parameters estimated by truncated least squares and GTLS ($\hat{\theta}$), with data set 1 the calibration period and data set 2 the validation period, are presented in Table 2. The predicted and measured temperatures of truncated least squares and GTLS are given in Figs. 3 and 4.

The same estimation, validation and cross-validation procedures were performed but now with the data-sets switched, *i.e.* the second data-set was used for estimation of $\hat{p}$ and $\hat{\theta}$. The results are given in Table 3 and Figs. 5 and 6.

GTLS has a good performance in each of the four cases and clearly outperforms the nonlinear least

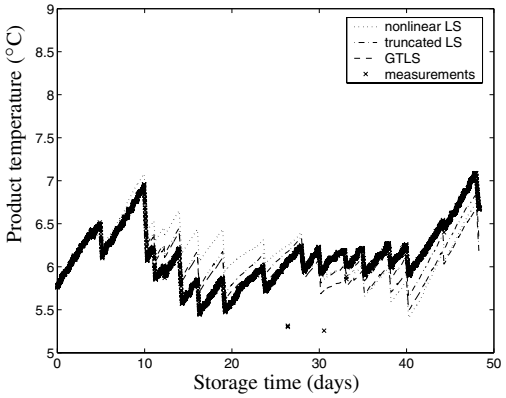

Fig. 4. Measured and predicted product temperatures in the validation period (data set 2) with parameters obtained by different calibration techniques.

Table 3. MSE of predictors with parameters estimated by nonlinear LS, truncated LS and GTLS in calibration (data set 2) and validation (data set 1) period.

	nonlinear LS	truncated LS	GTLS
calibration	0.097	0.035	0.017
validation	0.036	0.766	0.016

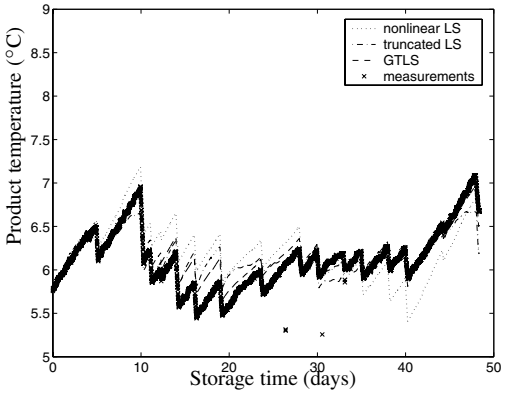

Fig. 5. Measured and predicted product temperatures in the calibration period (data set 2) with parameters obtained by different calibration techniques.

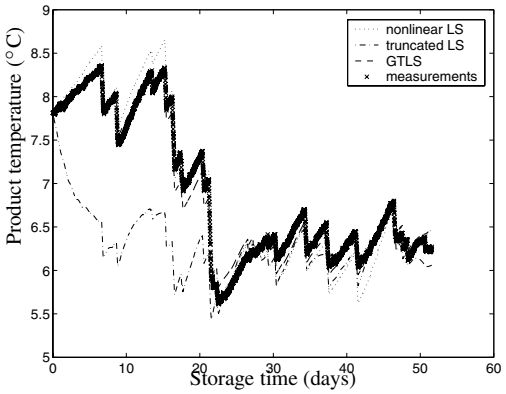

Fig. 6. Measured and predicted product temperatures in the validation period (data set 1) with parameters obtained by different calibration techniques.

squares estimates. The ordinary least squares estimate only has a good performance if the calibration dataset is informative enough. This can be clearly seen in the first part of figure 6 where the ordinary least squares estimate has a very poor performance. There appears to be some time variation in Figures 3 to 6, with estimated values above the actual values early on and below actual values late on. This is particularly the case for the nonlinear least squares estimates.

If a nonlinear model is linearly regressive reparameterized (from (2) to (4)) not only ordinary least squares or total least squares methods become available but a whole set of linear identification tools. As we focus on long-term predictive quality, a prediction error approach may be a good alternative to the presented least squares methods.

5. CONCLUSIONS

Reparameterization of nonlinear discrete-time models with polynomial quotient structure in Z and p towards linear regression can result in estimates that lead to good predictive quality. In a simulation environment with low noise signals GTLS outperforms ordinary least squares in predictive context. In a real world example it is shown that if the calibration data set is informative, ordinary least squares estimation generates good results in predictive context. For a less informative dataset, the ordinary least squares estimate behaves badly in the region where no information is available. GTLS generates good results in both cases.

ACKNOWLEDGEMENTS

We are indebted to Tolsma for providing the data. This research is part of the WIC2 project, which is carried out in co-operation with PRIVA BV, Tolsma, WNI and Agrotechnology and Food Innovations. It is financially supported with EET-subsidy from the Dutch Ministry of Finance, of Education, Culture and Science, and of Housing, Regional Development and Environment (project no. EETK01120).

REFERENCES

Doeswijk, T. G. and K. J. Keesman (2005). Parameter estimation and prediction of a nonlinear storage model: an algebraic approach. In: *Proceedings of the 2005 International Conference on Control & Automation*. IEEE. Budapest.

Godfrey, K (1983). *Compartmental models and their application*. Academic Press. London.

Golub, G. H. and C. F. Van Loan (1980). An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis* **17**(6), 883–893.

Holmberg, A. (1982). On the practical identifiability of microbial-growth models incorporating michaelis-menten type nonlinearities. *Mathematical Biosciences* **62**(1), 23–43.

Keesman, K. J., D. Peters and L. J. S. Lukasse (2003). Optimal climate control of a storage facility using local weather forecasts. *Control Engineering Practice* **11**(5), 505–516.

Ljung, L. (1987). *System Identification, Theory for the User*. Information and System Sciences. Prentice Hall, Inc.. Englewood Cliffs, New Jersey.

Norton, J. P. (1986). *An Introduction to Identification*. Academic Press. London.

Van Huffel, S. and J. Vandewalle (1989). Analysis and properties of the generalized total least squares problem $Ax \approx b$ when some or all columns in A are subject to error. *SIAM Journal on Matrix Analysis and Applications* **10**, 294–315.

Van Huffel, S. and J. Vandewalle (1991). *The Total Least Squares Problem: Computational Aspects and Analysis*. Vol. 9 of *Frontiers in Applied Mathematics*. SIAM. Philadelphia.

Van Huffel, S. and Lemmerling, P., Eds. (2002). *Total Least Squares and Error-in-Variables modeling: Analysis, Algorithms and Applications*. Kluwer Academic Publishers. Dordrecht.