ON THE ESTIMATION OF NON-STATIONARY FUNCTIONAL SERIES TARMA MODELS

Aggelos G. Poulimenos and Spilios. D. Fassois

Stochastic Mechanical Systems & Automation (SMSA) Laboratory
Department of Mechanical and Aeronautical Engineering
University of Patras, GR 26500 Patras, Greece
Tel/Fax: (++30) 2610 997405, e-mail: {poulimen,fassois}@mech.upatras.gr
Internet: www.mech.upatras.gr/∼sms

ABSTRACT

Maximum Likelihood (ML) and Multi-Stage Weighted Linear / Non-Linear Least Squares (MS-WLLS / MS-WNLS) estimation methods are formulated for non-stationary Functional Series Time-dependent ARMA (FS-TARMA) models. The methods' effectiveness, as well as their superiority over Ordinary Linear / Non-Linear Least Squares (OLLS / ONLS) type methods not accounting for innovations serial heteroscedasticity, are demonstrated via Monte Carlo experiments.

1. INTRODUCTION

Non-stationary stochastic signals, that is signals with time-dependent characteristics, are frequently encountered in engineering, and have been receiving increasing attention in recent years [1]. Examples include signals relating to seismic motion, speech, rotating machinery, and so on.

Parametric methods for the modelling and analysis of non-stationary stochastic signals complement their non-parametric counterparts and offer advantages such as representation parsimony, improved accuracy, resolution, and tracking [2]. A notable class of such methods is based upon *Functional Series Time-dependent AutoRegressive Moving Average (FS-TARMA)* models, which resemble their conventional ARMA counterparts with the important difference that their parameters and innovations variance are time-dependent, belonging to specific functional subspaces [1, 2, 3]. FS-TARMA methods are particularly attractive over alternative approaches, as they offer high parsimony, the capability of tracking "fast" or "slow" variations in the dynamics, as well as high achievable accuracy and resolution [2].

Functional Series TARMA model estimation has been, thus far, mainly based upon Ordinary Linear / Non-Linear Least Squares (OLLS / ONLS) type methods [1, 2]. The resulting estimators are, in the general case of time-dependent innovations variance (the *heteroscedastic* case), inefficient. On the other hand, Maximum Likelihood (ML) estimation leads to non-linear estimators of increased complexity even in the pure AutoRegressive (FS-TAR) case [4].

The *goal* of the present study is the formulation of Maximum Likelihood (ML) and Multi-Stage Weighted Linear / Non-Linear Least Squares (MS-WLLS / MS-WNLS) estimators for FS-TAR/TARMA models, as well as their testing and comparison to Ordinary Linear / Non-Linear Least Squares (OLLS / ONLS) estimators via Monte Carlo experiments.

2. FUNCTIONAL SERIES TARMA MODELS

FS-TARMA models constitute conceptual extensions of their conventional (stationary) ARMA counterparts, in that their parameters are explicit functions of time by belonging to functional subspaces spanned by selected deterministic functions (*basis functions*). An FS-TARMA$(n_a,n_c)_{[p_a,p_c,p_s]}$ model, with n_a, n_c designating its AutoRegressive (AR) and Moving Average (MA) orders, respectively, and p_a, p_c, p_s its AR, MA and innovations standard deviation functional basis dimensionalities, respectively, is of the form [1, 2]:

$$\underbrace{x[t]+\sum_{i=1}^{n_a} a_i[t]\cdot x[t-i]}_{\text{AR-part}} = \underbrace{e[t]+\sum_{i=1}^{n_c} c_i[t]\cdot e[t-i]}_{\text{MA-part}} \qquad (1)$$

with t designating normalized discrete time, $x[t]$ the non-stationary signal modelled, and $e[t]$ an uncorrelated (white) innovations (residual) sequence with zero mean and time-dependent variance $\sigma_e^2[t]$. $a_i[t]$ and $c_i[t]$ stand for the i-th time-dependent AR and MA parameter, respectively, which, along with the residual standard deviation, belong to their respective functional subspaces:

$$\begin{aligned}
\mathscr{F}_{AR} &\triangleq \{G_{b_a(1)}[t],\, G_{b_a(2)}[t],\, \ldots,\, G_{b_a(p_a)}[t]\}\\
\mathscr{F}_{MA} &\triangleq \{G_{b_c(1)}[t],\, G_{b_c(2)}[t],\, \ldots,\, G_{b_c(p_c)}[t]\}\\
\mathscr{F}_{\sigma_e} &\triangleq \{G_{b_s(1)}[t],\, G_{b_s(2)}[t],\, \ldots,\, G_{b_s(p_s)}[t]\}
\end{aligned}$$

In these expressions the indices $b_a(j)$ $(j=1,\ldots,p_a)$, $b_c(j)$ $(j=1,\ldots,p_c)$ and $b_s(j)$ $(j=1,\ldots,p_s)$ designate the functions (from a properly ordered set, such as Chebyshev or other polynomials) that are included in each basis. The time-dependent AR/MA parameters and the residual standard deviation may be thus expressed as:

$$a_i[t] \triangleq \sum_{j=1}^{p_a} a_{i,j}\cdot G_{b_a(j)}[t], \qquad c_i[t] \triangleq \sum_{j=1}^{p_c} c_{i,j}\cdot G_{b_c(j)}[t]$$
$$\sigma_e[t] \triangleq \sum_{j=1}^{p_s} s_j\cdot G_{b_s(j)}[t]$$

with $a_{i,j}$, $c_{i,j}$ and s_j designating the corresponding *coefficients of projection*. The model is thus parameterized in terms of the projection coefficients $a_{i,j}$, $c_{i,j}$, s_j, and the model structure ($\mathscr{M}$) parameters (the model orders n_a, n_c, and the functional subspace indices).

3. MAXIMUM LIKELIHOOD ESTIMATION OF FS-TARMA MODELS

Maximum Likelihood (ML) estimation of the model projection coefficient vector [1]:

$$\boldsymbol{\theta} \stackrel{\Delta}{=} [\ \boldsymbol{\vartheta}^T \mid \mathbf{s}^T\ ]^T$$

$$\boldsymbol{\vartheta} \stackrel{\Delta}{=} [\ a_{1,1}\ \ldots\ a_{n_a,p_a} \mid c_{1,1}\ \ldots\ c_{n_c,p_c}\ ]^T, \quad \mathbf{s} \stackrel{\Delta}{=} [s_1 \ldots s_{p_s}]^T$$

is presently considered based upon available signal samples $x^N = \{x[1] \ldots x[N]\}$ and a given model structure $\mathscr{M}$.

The ML estimator is defined as the estimator that maximizes the likelihood of the unknown vector $\boldsymbol{\theta}$ given the observations x^N:

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} \ln \mathscr{L}(\boldsymbol{\theta}|x^N)$$

$$\mathscr{L}(\boldsymbol{\theta}|x^N) \stackrel{\Delta}{=} f(x^N|\boldsymbol{\theta}) = f(e^N|\boldsymbol{\theta})$$

with $\mathscr{L}(\cdot)$ designating the likelihood function, $f(\cdot)$ probability density function, and $e^N = \{e[1], \ldots, e[N]\}$. Assuming Gaussian observations, the conditional (upon the initial conditions) form of the log-likelihood is:

$$\ln f(e^N|\boldsymbol{\theta}) = \ln \prod_{t=1}^{N} f(e[t]|\boldsymbol{\theta}) = \sum_{t=1}^{N} \ln f(e[t]|\boldsymbol{\theta}) =$$

$$= \sum_{t=1}^{N} \ln \left((2\pi\sigma_e^2[t,\mathbf{s}])^{-1/2} \cdot \exp\left\{ \frac{-e^2[t,\boldsymbol{\vartheta}]}{2\sigma_e^2[t,\mathbf{s}]} \right\} \right) =$$

$$= -\frac{N}{2} \cdot \ln 2\pi - \frac{1}{2} \cdot \sum_{t=1}^{N} \left(\ln(\mathbf{g}_s^T[t] \cdot \mathbf{s})^2 + \frac{e^2[t,\boldsymbol{\vartheta}]}{(\mathbf{g}_s^T[t] \cdot \mathbf{s})^2} \right) \quad (2a)$$

$$\mathbf{g}_s[t] \stackrel{\Delta}{=} \left[G_{b_s(1)}[t]\ G_{b_s(2)}[t]\ \ldots\ G_{b_s(p_s)}[t] \right]^T \quad (2b)$$

with $e[t,\boldsymbol{\vartheta}]$ designating the residual corresponding to the $\boldsymbol{\vartheta}$ parameter vector and obtained via Eq. (1).

Maximization of the log-likelihood function of Eq. (2a) constitutes a *non-linear* optimization problem that has to be handled via iterative techniques. This is due to the non-quadratic dependence of the log-likelihood function upon the MA coefficients of projection (TARMA case) and the innovations standard deviation coefficients of projection (TAR/TARMA cases).

It is, nevertheless, well known that iterative non-linear optimization techniques are, in particular in the present case, amenable to acute wrong convergence problems typically due to the existence of several local maxima [2, 5]. For this reason quite accurate parameter estimates are normally required for starting the optimization. Such estimates may be provided by Multi-Stage Weighted Linear / Non-Linear Least Squares (MS-WLLS / MS-WNLS) based estimation methods which are discussed next. Notice that (depending upon the desired accuracy) the MS-WLLS / MS-WNLS methods may be used either as stand-alone (suboptimal) schemes or as essential parts of (optimal) ML estimation.

[1] Bold face symbols designate (column) vector quantities.

4. MULTI-STAGE WEIGHTED LEAST SQUARES ESTIMATION OF FS-TARMA MODELS

Multi-Stage methods aim at the (suboptimal) decoupling of the complete estimation problem into a sequence of simpler subproblems. This presently amounts to decoupling the estimation problem into AR/MA coefficients of projection estimation and innovations standard deviation coefficients of projection estimation. In the mixed FS-TARMA case the AR/MA coefficients of projection estimation may be further decoupled into separate AR and MA estimation procedures. These decouplings facilitate estimation, as they permit the use of Weighted Linear / Non-Linear Least Squares (in the non-linear case the benefit being the reduction in the optimization space dimensionality) and other techniques (such as convolution and deconvolution operations) at the price of statistical inefficiency. The resulting deterioration in the achievable accuracy may be, nevertheless, small.

The first of the above mentioned decouplings is motivated by the fact that, given the true innovations standard deviation coefficients of projection, say $\mathbf{s}^\circ$, the log-likelihood simplifies to:

$$\ln f(e^N|\boldsymbol{\vartheta}) = -\left(\frac{1}{2} \cdot \sum_{t=1}^{N} \frac{e^2[t,\boldsymbol{\vartheta}]}{(\mathbf{g}_s^T[t] \cdot \mathbf{s}^\circ)^2} \right) + \text{const.} \quad (3)$$

In the pure FS-TAR case $\boldsymbol{\vartheta} \stackrel{\Delta}{=} \mathbf{a}$ (the AR coefficients of projection vector) and the dependence of $e[t,\boldsymbol{\vartheta}]$ upon it is linear. This implies that maximization of the above likelihood is a quadratic problem, leading to a Weighted Linear Least Squares (WLLS) estimator. In the mixed FS-TARMA case a subsequent decoupling [as in the Polynomial-Algebraic (P-A) or the Two Stage Least Squares (2SLS) methods [2]] may be used in order to convert the problem into a sequence of linear subproblems. The obtained estimates may be optionally refined via weighted non-linear least squares.

4.1 MS-WLLS Estimation of FS-TAR Models

Stage 1. Initial AR Coefficients of Projection Estimation.

Since the vector $\mathbf{s}$ is not a-priori available, initial estimation of $\boldsymbol{\vartheta} \equiv \mathbf{a}$ is achieved via the Ordinary Linear Least Squares (OLLS) estimator:

$$\hat{\boldsymbol{\vartheta}}^{in} \equiv \hat{\mathbf{a}}^{in} = \arg\min_{\boldsymbol{\vartheta}} \sum_{t=1}^{N} \left(x[t] - \boldsymbol{\phi}^T[t] \cdot \boldsymbol{\vartheta} \right)^2 \implies$$

$$\implies \hat{\boldsymbol{\vartheta}}^{in} \equiv \hat{\mathbf{a}}^{in} = \left(\sum_{t=1}^{N} \boldsymbol{\phi}[t] \cdot \boldsymbol{\phi}^T[t] \right)^{-1} \cdot \left(\sum_{t=1}^{N} \boldsymbol{\phi}[t] \cdot x[t] \right) \quad (4)$$

$$\boldsymbol{\phi}[t] \stackrel{\Delta}{=} \left[-G_{b_a(1)}[t] \cdot x[t-1]\ \ldots\ -G_{b_a(p_a)}[t] \cdot x[t-n_a] \right]^T$$

Stage 2. Residual Standard Deviation Estimation.

The residual series $e[t,\hat{\boldsymbol{\vartheta}}^{in}]$ is obtained via the model expression of Eq. (1) using the estimate $\hat{\boldsymbol{\vartheta}}^{in} \equiv \hat{\mathbf{a}}^{in}$. Estimation of the residual standard deviation coefficients of projection (vector $\mathbf{s}$) may be then obtained as [6]:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \frac{1}{N} \sum_{t=1}^{N} \left(|e[t,\hat{\boldsymbol{\vartheta}}^{in}]| - \sqrt{\frac{2}{\pi}} \cdot \mathbf{g}_s^T[t] \cdot \mathbf{s} \right)^2 \implies$$

$$\Longrightarrow \hat{\mathrm{s}} = \sqrt{\frac{\pi}{2}} \cdot \left(\sum_{t=1}^{N} \mathbf{g}_s[t] \cdot \mathbf{g}_s^T[t] \right)^{-1} \cdot \left(\sum_{t=1}^{N} \mathbf{g}_s[t] \cdot |e[t, \hat{\vartheta}^{in}]| \right) \quad (5)$$

For increased accuracy these estimates may be refined by maximizing the log-likelihood of s given $e[t, \hat{\vartheta}^{in}]$ (now treated as observations), that is:

$$\hat{\mathrm{s}} = \arg\max_{\mathrm{s}} \left\{ -\frac{1}{2} \sum_{t=1}^{N} \left(\ln(\mathbf{g}_s^T[t] \cdot \mathrm{s})^2 + \frac{e^2[t, \hat{\vartheta}^{in}]}{(\mathbf{g}_s^T[t] \cdot \mathrm{s})^2} \right) \right\} \quad (6)$$

Estimation of s based upon this expression constitutes a non-linear optimization problem that may be tackled via proper (iterative) techniques.

Stage 3. Final AR Coefficients of Projection Estimation.

Once $\hat{\mathrm{s}}$ is available, final estimation of the AR coefficients of projection may be achieved by maximizing the log-likelihood function of Eq. (3). This leads to the Weighted Linear Least Squares (WLLS) estimator:

$$\hat{\vartheta} \equiv \hat{\mathbf{a}} = \arg\min_{\vartheta} \sum_{t=1}^{N} \frac{\left(x[t] - \phi^T[t] \cdot \vartheta\right)^2}{(\mathbf{g}_s^T[t] \cdot \hat{\mathrm{s}})^2} \Longrightarrow$$

$$\Longrightarrow \hat{\vartheta} \equiv \hat{\mathbf{a}} = \left(\sum_{t=1}^{N} \frac{\phi[t] \cdot \phi^T[t]}{(\mathbf{g}^T[t] \cdot \hat{\mathrm{s}})^2} \right)^{-1} \cdot \left(\sum_{t=1}^{N} \frac{\phi[t] \cdot x[t]}{(\mathbf{g}^T[t] \cdot \hat{\mathrm{s}})^2} \right) \quad (7)$$

4.2 MS-WLLS and MS-WNLS Estimation of FS-TARMA Models

Stage 1. Initial AR/MA Coefficients of Projection Estimation.

Initial estimation of the AR/MA coefficients of projection (vector ϑ) may be achieved via the estimator:

$$\hat{\vartheta}^{in} = \arg\min_{\vartheta} \sum_{t=1}^{N} e^2[t, \vartheta] \quad (8)$$

Since the cost function is non-quadratic in terms of the MA coefficients of projection, the problem may be (suboptimally) converted into a sequence of linear subproblems via the Polynomial-Algebraic (P-A) or the Two Stage Least Squares (2SLS) methods [2] *(MS-WLLS version)*. For increased accuracy, these estimates may be refined by the Non-Linear Least Squares (NLS) estimator of Eq. (8), implemented via iterative optimization techniques *(MS-WNLS version)*.

Stage 2. Residual Standard Deviation Estimation.

The residual series $e[t, \hat{\vartheta}^{in}]$ is obtained via the model expression of Eq. (1) using the estimate $\hat{\vartheta}$. Estimation of the residual standard deviation coefficients of projection (vector s) is then achieved as in Stage 2 of the FS-TAR case.

Stage 3. Final AR/MA Coefficients of Projection Estimation.

Once $\hat{\mathrm{s}}$ is available, the final estimation of the AR/MA coefficients of projection may be based upon maximization of the log-likelihood function of Eq. (3). Since this criterion is non-quadratic in terms of the MA coefficients of projection, the problem may be (suboptimally) converted into a sequence of linear subproblems via properly adjusted (weighted) versions of the Polynomial-Algebraic (P-A) or the Two Stage Least Squares (2SLS) methods [2] *(MS-WLLS version)*. For increased accuracy, these estimates may be refined by the Weighted Non-Linear Least Squares (WNLS) estimator implied by Eq. (3), implemented via iterative optimization techniques *(MS-WNLS version)*.

Table 1: FS-TAR(2) ML estimation results (500 runs).

Projection coefficient	Theoretical value	ML estimate (mean ± std deviation)
$a_{1,1}$	−0.1269	−0.1264 ± 0.0087
$a_{1,2}$	−0.2050	0.2043 ± 0.0078
$a_{1,3}$	0.2058	0.2058 ± 0.0075
$a_{2,1}$	0.7401	0.7390 ± 0.0083
$a_{2,2}$	0.0180	0.0176 ± 0.0082
$a_{2,3}$	−0.0180	−0.0176 ± 0.0072
s_1	1.5000	1.4986 ± 0.0149
s_2	−0.6000	−0.5994 ± 0.0169
s_3	0.4000	0.3990 ± 0.0135

Table 2: FS-TARMA(2,2) ML estimation results (500 runs).

Projection coefficient	Theoretical value	ML estimate (mean ± std deviation)
$a_{1,1}$	−0.3611	−0.3603 ± 0.0118
$a_{1,2}$	0.0604	0.0605 ± 0.0071
$a_{1,3}$	0.3129	0.3128 ± 0.0087
$a_{2,1}$	0.7610	0.7603 ± 0.0115
$a_{2,2}$	−0.0048	−0.0052 ± 0.0070
$a_{2,3}$	−0.0283	−0.0281 ± 0.0078
$c_{1,1}$	−0.1863	−0.1863 ± 0.0174
$c_{1,2}$	−0.0699	−0.0702 ± 0.0122
$c_{1,3}$	0.0937	0.0940 ± 0.0147
$c_{2,1}$	0.0503	0.0508 ± 0.0174
$c_{2,2}$	0.0141	0.0144 ± 0.0120
$c_{2,3}$	−0.0191	−0.0192 ± 0.0141
s_1	0.9000	0.8989 ± 0.0088
s_2	0.4000	0.3995 ± 0.0070
s_3	−0.2000	−0.1999 ± 0.0059

5. MONTE CARLO EXPERIMENTS

The effectiveness of FS-TAR/TARMA Maximum Likelihood (ML) estimation, based upon initialization via the Multi-Stage Weighted Linear / Non-Linear Least Squares (MS-WLLS / MS-WNLS) methods, is examined via two Monte Carlo experiments (MS-WLLS is used in the FS-TAR case; MS-WNLS combined with 2SLS estimation, maximum of 10 2SLS iterations and inverse function order of 12, is used in the FS-TARMA case).

Non-stationary signals generated by an FS-TAR$(2)_{[3,3]}$ model with Chebyshev II functional subspaces ($\mathscr{F}_{AR} = \mathscr{F}_{\sigma_e} = \{G_0[t], G_1[t], G_2[t]\}$) are used in the first experiment, and signals generated by an FS-TARMA$(2,2)_{[3,3,3]}$ model, also with Chebyshev II functional subspaces ($\mathscr{F}_{AR} = \{G_0[t], G_2[t], G_3[t]\}$, $\mathscr{F}_{MA} = \mathscr{F}_{\sigma_e} = \{G_0[t], G_1[t], G_3[t]\}$), are used in the second experiment. The true model coefficients of projection are shown in Tables 1 (FS-TAR case) and 2 (FS-TARMA case).

Each experiment consists of 500 runs, with each run based upon an $N = 6000$ sample-long signal realization. The Maximum Likelihood (ML) FS-TAR/TARMA coefficient of

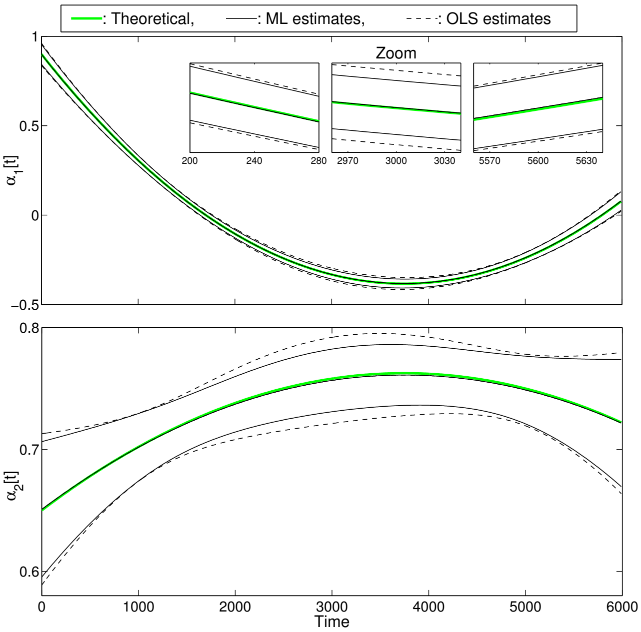

Figure 1: FS-TAR(2) experiment: The theoretical, ML-estimated, and OLLS-estimated AR parameter trajectories (mean estimates $\pm$ two standard deviations; 500 runs).

projection estimation results are summarized in Tables 1 and 2, respectively, and are shown to be in excellent agreement with their theoretical counterparts, being characterized by reasonably small standard deviations.

The theoretical and ML-estimated (mean $\pm$ two standard deviations) AR/MA parameter trajectories are, as functions of time, depicted in Figures 1 and 2 for the FS-TAR(2) and FS-TARMA$(2,2)$ cases, respectively. In the first figure the corresponding Ordinary Linear Least Squares (OLLS) estimated trajectories [estimator of Eq. (4)] are also shown (mean $\pm$ two standard deviations) for purposes of comparison. Similarly, the corresponding Ordinary Non-Linear Least Squares (ONLS; 2SLS based initialization) estimated trajectories [estimator of Eq. (8)] are shown in Figure 2. The basic difference between the OLLS/ONLS and the ML estimators is that the former do *not* account for the serially heteroscedastic (time-varying variance) nature of the innovations. The two types of estimates are, in both experiments, similar, although the ML estimates expectedly exhibit somewhat improved standard deviations.

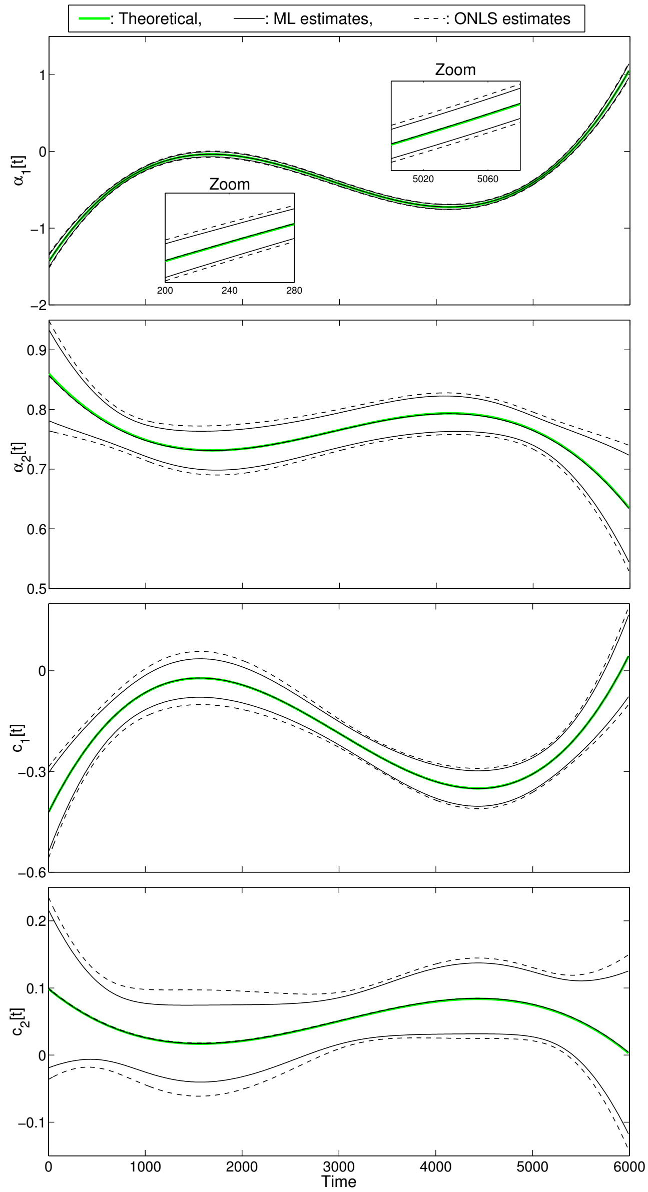

Figure 2: FS-TARMA(2,2) experiment: The theoretical, ML-estimated, and ONLS-estimated AR/MA parameter trajectories (mean estimates $\pm$ two standard deviations; 500 runs).

6. CONCLUSIONS

In this study Maximum Likelihood (ML) and Multi-Stage Weighted Linear / Non-Linear Least Squares (MS-WLLS / MS-WNLS) estimators were formulated for FS-TAR/TARMA models. The effectiveness of the ML estimators and their superiority over their Ordinary Linear / Non-Linear Least Squares (OLLS / ONLS) counterparts were demonstrated via Monte Carlo experiments.

REFERENCES

[1] M. Niedzwiecki, *Identification of Time-Varying Processes*. John Wiley & sons, 2000.

[2] A.G. Poulimenos, and S.D. Fassois, “Non-stationary random vibration modelling and analysis via Functional Series TARMAX models,” in *Proc. ISMA Intern. Conf. Noise and Vibration Engineering*, Leuven, 2004.

[3] F. Kozin, and F. Nakajima, “The order determination problem for linear time-varying AR models,” *IEEE Trans. Automatic Control*, vol. 25, pp. 250–257, 1980.

[4] R. Dahlhaus, “Fitting time series models to non-stationary processes,” *The Annals of Statistics*, vol. 25, pp. 1–37, 1997.

[5] A.G. Poulimenos, and S.D. Fassois, “On the efficient estimation of non-stationary Functional Series TARMA models,” *Under preparation*.

[6] Y. Grenier, “Time-dependent ARMA modeling of non-stationary signals,” *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 31, pp. 899–911, 1983.