Chapter 73
ARM Processes and Their Modeling and Forecasting Methodology

Benjamin Melamed

Abstract The class of *ARM (Autoregressive Modular)* processes is a class of stochastic processes, defined by a nonlinear autoregressive scheme with modulo-1 reduction and additional transformations. ARM processes constitute a versatile class designed to produce high-fidelity models from stationary empirical time series by fitting a strong statistical signature consisting of the empirical marginal distribution (histogram) and the empirical autocorrelation function. More specifically, fitted ARM processes guarantee the matching of arbitrary empirical distributions, and simultaneously permit the approximation of the leading empirical autocorrelations. Additionally, simulated sample paths of ARM models often resemble the data to which they were fitted. Thus, ARM processes aim to provide both a quantitative and qualitative fit to empirical data. Fitted ARM models can be used in two ways: (1) to generate realistic-looking Monte Carlo sample paths (e.g., financial scenarios), and (2) to forecast via point estimates as well as confidence intervals.

This chapter starts with a high-level discussion of stochastic model fitting and explains the fitting approach that motivates the introduction of ARM processes. It then continues with a review of ARM processes and their fundamental properties, including their construction, transition structure, and autocorrelation structure. Next, the chapter proceeds to outline the ARM modeling methodology by describing the key steps in fitting an ARM model to empirical data. It then describes in some detail the ARM forecasting methodology, and the computation of the conditional expectations that serve as point estimates (forecasts of future values) and their underlying conditional distributions from which confidence intervals are constructed for the point estimates. Finally, the chapter concludes with an illustration of the efficacy of the ARM modeling and forecasting methodologies through an example utilizing a sample from the S&P 500 Index.

B. Melamed (✉)
Department of Supply Chain Management and Marketing Sciences,
Rutgers Business School – Newark and New Brunswick,
Rutgers University, New Brunswick, NJ, USA
e-mail: melamed@rbs.rutgers.edu

Keywords ARM processes • Autoregressive modular processes • ARM modeling methodology • ARM forecasting methodology • Financial applications

73.1 Introduction

The enterprise of modeling aims to *fit* a model from a parametric family to empirical data, subject to prescribed fitting criteria. From a high-level vantage point, fitting criteria formulate some measure of "deviation" of any candidate model from the empirical data, and modeling proceeds as a search for a model that minimizes that "deviation" measure. For example, regression-type modeling utilizes fitting criteria such as least squares, maximum likelihood, and others.

A *statistical signature* (*signature*, for short) is a set of empirical or theoretic statistics: the former is computed from empirical data and the latter is derived from a theoretic model. Here, we shall be concerned with time series modeling under stationary conditions, so the signatures of interest will be derived from stationary empirical time series and stationary time-series models. Accordingly, members of a signature might include first-order statistics (e.g., the common mean, variance and/or higher moments, or the marginal distribution), second-order statistics (e.g., the autocorrelation function, or equivalently, the spectral density), or even higher-order statistics. Clearly, statistical signatures can be partially ordered by strength under set inclusion in a natural way as follows: signature A is stronger (dominates) signature B, if A contains B.

In *signature-oriented* modeling, one simply utilizes fitting criteria based on a statistical signature. In other words, for a prescribed signature, one searches for a model whose signature fits or approximates its empirical counterpart. It is reasonable to expect that the predictive power of models (referred to as *model fidelity*) would generally increase as the fitting signature becomes stronger; stated more simply, the more empirical statistics are fitted by a model, the better the model. The need for high-fidelity models of stochastic sequences is self-evident in systems containing random

C.-F. Lee et al. (eds.), *Handbook of Quantitative Finance and Risk Management*,
DOI 10.1007/978-0-387-77117-5_73,

phenomena or components, such as temporal dependence in the context of queueing systems (Livny et al. 1993; Altiok and Melamed 2001). Examples include bursty traffic in the Next Generation Internet (especially compressed video – a prodigious consumer of bandwidth), bursty arrivals of demand at an inventory system, reliability (breakdown and repair) of components with non-independent uptimes and downtimes, interest rate evolution, and unfolding enemy arrivals (engagements) in a battlefield simulation.

This chapter reviews a class of high fidelity and versatile models, called *ARM* (*Autoregressive Modular*) processes. It then proceeds to describe the ARM modeling and forecasting methodology. All variants of ARM processes share a signature-oriented modeling philosophy, where the underlying empirical time series is stationary in the sense that the underlying probability law is assumed to have at least time invariant 2-dimensional joint distributions. Specifically, ARM processes aim to construct high-fidelity models from empirical data, which can capture a strong statistical signature of the empirical data by *simultaneously* fitting the *empirical marginal distribution* and the *empirical leading (low-lag) autocorrelations*. Accordingly, a prospective model is declared a good fit to an empirical time series, provided it satisfies the following requirements:

(a) The marginal distribution of the model fits exactly its empirical counterpart (histogram).
(b) The leading autocorrelations of the model approximate their empirical counterparts (the number of lags is left to the modeler and often determined by the availability of data).
(c) The model gives rise to Monte Carlo simulated paths that are visually "similar" to ("resemble") the empirical records.

Requirements (a) and (b) constitute crisp quantitative goodness-of-fit criteria (see also Cario and Nelson 1996). In contrast, the notions of sample path "similarity" or "resemblance" in requirement (c) are heuristic and cannot be defined with mathematical crispness; note that "resemblance" does not mean that the sample paths are identical point for point, but rather that human observers perceive them as "similar" in appearance. It should be pointed out that these notions are intuitively clear and that modelers routinely attempt to validate their models by visual inspections for "path similarity." Experience suggests that ARM processes can be fitted to arbitrary marginal distributions, a wide variety of autocorrelation functions (monotone, oscillating, alternating, etc.), and a broad range of sample path behaviors (non-directional, cyclical, etc.)

The following notation will be used throughout. The Laplace transform of a function $f(x)$ is defined by $\tilde{f}(s) = \int_{-\infty}^{\infty} f(x)\, e^{-sx}\, dx$ (tilde denotes the Laplace transform operator). The cumulative distribution function (cdf) of a random variable Z is denoted by F_Z and its probability density function (pfd) is denoted by f_Z; the corresponding mean and standard deviation are denoted by μ_Z and σ_Z, respectively. Similarly, the marginal density of a stationary random sequence $\{Z_n\}$ is denoted by f_Z. The notation $Z \sim F$ means that the random variable Z has the cdf F. The autocorrelation function of a real-valued random sequence $\{Z_n\}$ with finite second moments is

$$\rho_Z(k,\tau) \;=\; \frac{\mathrm{E}[Z_k\, Z_{k+\tau}] \;-\; \mathrm{E}[Z_k]\,\mathrm{E}[Z_{k+\tau}]}{\sigma_{Z_k}\,\sigma_{Z_{k+\tau}}} \tag{73.1}$$

When $\rho_Z(k,\tau)$ does not depend on k in Equation (73.1), but only on the lag τ, we denote the autocorrelation function by $\rho_Z(\tau)$. The conditional density of a random variable Y given a random variable X is denoted by

$$f_{Y|X}(y|x) \;=\; \frac{\partial}{\partial y}\, P\{Y \;\le\; y \,|\, X \;=\; x\} \tag{73.2}$$

The indicator function of a set A is denoted either by $1_A\,(x)$ or $1\{x \in A\}$, while symbols decorated by a hat always denote a statistical estimator. The imaginary number $\sqrt{-1}$ is denoted by i, and the real-part operator by Re. Finally, the modulo-1 (fractional part) operator, $\langle\rangle$, is defined by $\langle x\rangle = x - \max\{\text{integer}\, n : n \le x\}$.

73.2 Overview of ARM Processes

This section provides an overview of ARM processes – a class of autoregressive stochastic processes, with modulo-1 reduction and additional transformations (Melamed 1999). It generalizes the class of *TES* (*Transform-Expand-Sample*) processes (Melamed 1991; Jagerman and Melamed 1992a, 1992b, 1994, 1995; Melamed 1993, 1997), and it further gives rise to a discretized version that subsumes the class of *QTES* (*Quantized TES*) processes (Melamed et al. 1996; Klaoudatos et al. 1999), a discretized variant of TES. However, while TES and QTES processes admit only iid (independent identically distributed) so-called *innovation sequences* (see below), ARM processes constitute a considerable generalization in that they admit dependent innovation sequences as well.

73.2.1 Background ARM Processes

We define two variants of ARM process classes, denoted by ARM$^+$ and ARM$^-$, respectively, and henceforth we append a plus or minus superscript to objects associated with these classes, but omit the superscript if the object is common to both classes. Accordingly, let $U_0 \sim \mathrm{Unif}(0,\ 1)$

be a random variable, uniformly distributed on the interval [0,1). An *innovation process* is any real-valued stochastic sequence, $\{V_n\}_{n=0}^{\infty}$, where the sequence $\{V_n\}$ is independent of U_0. All random variables are defined over some common probability space (Ω, Σ, P).

We next define two stochastic sequences, $\{U_n^+\}_{n=0}^{\infty}$ and $\{U_n^-\}_{n=0}^{\infty}$, to be referred to as *background ARM processes* due to their auxiliary nature. These are given by the recursive equations

$$U_n^+ = \begin{cases} \langle U_0 + V_0 \rangle, & n = 0 \\ \langle U_{n-1}^+ + V_n \rangle, & n > 0 \end{cases} \tag{73.3}$$

$$U_N^- = \begin{cases} U_n^+, & n \text{ even} \\ 1 - U_n^+, & n \text{ odd} \end{cases} \tag{73.4}$$

We remark that the definitions above have a simple geometric interpretation: the modulo-1 arithmetic renders background ARM processes random walks on the unit circle (a circle of unit circumference), where each point on the circle corresponds to a fraction in the range [0,1) (see Diaconis 1988). The innovation random variables correspond to step sizes of the random walk on the unit circle. In fact, the properties of modulo-1 arithmetic imply the following alternative representation for Equation 73.3,

$$U_n^+ = \langle U_0 + S_{0,n} \rangle = \langle U_0 + \langle S_{0,n} \rangle \rangle, \tag{73.5}$$

where the $S_{m,n}$ are partial sums of innovation random variables, given by

$$S_{m,n} = \begin{cases} 0, & \text{if } m > n \\ \sum_{j=m}^{n} V_j, & \text{if } m \le n \end{cases} \tag{73.6}$$

The probability law of the innovation process determines whether or not the background process is directional. In particular, note that a non-symmetrical innovation density imparts a drift to the background process, resulting in a cyclical process around the unit circle. See Melamed (1999) for details.

73.2.2 Foreground ARM Processes

Target ARM processes are transformed background ARM processes, and are consequently referred to as *foreground ARM processes*. These are of the general forms

$$X_n^+ = D(U_n^+) \tag{73.7}$$

or

$$X_n^- = D(U_n^-) \tag{73.8}$$

where D is a measurable transformation from the interval [0,1) to the real line, called a *distortion*. We assume throughout that $\int_0^1 D^2(x)dx < \infty$. In the sequel, we shall drop the superscript of background or foreground ARM processes, when the distinction is immaterial.

Two fundamental properties of ARM processes facilitate ARM modeling (Melamed 1999):

- The marginal distribution of *all* background ARM processes is *uniform* on the interval [0,1), *regardless* of the innovation sequence selected!
- If the innovation sequence is stationary, so are the corresponding marginal distributions of background and foreground ARM processes.

Consequently, we gain broad modeling flexibility as follows:

- Any marginal distribution can be easily fitted to a foreground ARM process, by appeal to the *Inverse Transform* method: If $U \sim \text{Unif}(0, 1)$ and F is any (cumulative) distribution function, then $F^{-1}(U) \sim F$ (see, e.g., Bratley et al. 1987; Law and Kelton 1991).
- Since this is true *regardless* of the innovation process selected, we can, in principle, attempt the fitting of both the empirical (marginal) distribution *and* empirical autocorrelation function *simultaneously*. The two fittings are decoupled: the first is guaranteed an exact match via the Inversion Transform method, and the second is approximated by a suitable choice of an innovation process.

73.2.3 Transition Functions of Background ARM Processes

The transition densities of background ARM processes are given separately for ARM$^+$ and ARM$^-$ processes as follows. For any background ARM$^+$ process $\{U_n^+\}$ and $0 \le u, v < 1$, $\tau \ge 1$,

$$\begin{aligned} f_{U_{k+\tau}^+ | U_k^+}(v|u) &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\; e^{i\,2\,\pi\,\nu\,(v-u)} \\ &= 1 + 2 \sum_{\nu=1}^{\infty} \text{Re}\left[\tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\; e^{i\,2\,\pi\,\nu\,(v-u)} \right] \end{aligned} \tag{73.9}$$

and for any background ARM$^-$ process $\{U_n^-\}$ and $0 \le u,\ v < 1,\ \tau \ge 1,$

$$f_{U_{k+\tau}^-|U_k^-}(v|u) = \begin{cases} \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\ e^{i\,2\,\pi\,\nu\,(v-u)}, & k \text{ even},\ \tau \text{ even} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\ e^{i\,2\,\pi\,\nu\,(-v-u)}, & k \text{ even},\ \tau \text{ odd} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\ e^{i\,2\,\pi\,\nu\,(-v+u)}, & k \text{ odd},\ \tau \text{ even} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\ e^{i\,2\,\pi\,\nu\,(v+u)}, & k \text{ odd},\ \tau \text{ odd} \end{cases} \tag{73.10}$$

Equation (73.10) can be further simplified analogously to Equation (73.9).

73.2.4 Autocorrelation Functions of Foreground ARM Processes

Let $\{X_n\}_{n=0}^{\infty}$ be a foreground ARM process of the form of Equations (73.7) or (73.8) with a finite positive variance and autocorrelation function $\rho_X(k,\tau)$. Then for any foreground ARM$^+$ process, $\{X_n^+\}$, and $\tau \ge 1$,

$$\begin{aligned} \rho_X^+(k,\tau) &= \frac{1}{\sigma_X^2} \sum_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\,|\tilde{D}(i\,2\,\pi\,\nu)|^2 \\ &= \frac{2}{\sigma_X^2} \sum_{\nu=1}^{\infty} \mathrm{Re}\left[\tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\ |\tilde{D}(i\,2\,\pi\,\nu)|^2\right] \end{aligned} \tag{73.11}$$

and for any foreground ARM$^-$ process, $\{X_n^-\}$, and $\tau \ge 1$,

$$\rho_X^-(k,\tau) = \begin{cases} \frac{1}{\sigma_X^2} \sum\limits_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\,|\tilde{D}(i\,2\,\pi\,\nu)|^2, & k \text{ even},\ \tau \text{ even} \\ \frac{1}{\sigma_X^2} \sum\limits_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\ \tilde{D}^2(i\,2\,\pi\,\nu), & k \text{ even},\ \tau \text{ odd} \\ \frac{1}{\sigma_X^2} \sum\limits_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\,|\tilde{D}(i\,2\,\pi\,\nu)|^2, & k \text{ odd},\ \tau \text{ even} \\ \frac{1}{\sigma_X^2} \sum\limits_{\substack{\nu=-\infty \\ \nu \neq 0}}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\ \tilde{D}^2(-i\,2\,\pi\,\nu), & k \text{ odd},\ \tau \text{ odd} \end{cases} \tag{73.12}$$

Equation (73.12) can be further simplified analogously to Equation (73.11).

73.2.5 Empirical Distributions

Practical ARM modeling focuses on distortions based on inverse distribution functions via an application of the Inverse Transform (Bratley et al. 1987). For a realization of an empirical sequence $\{Y_n\}_{n=0}^{N}$, an empirical density is often estimated by a histogram, which gives the observed relative frequencies of various ranges via a step function of the form

$$h_Y(y) = \sum_{j=1}^{J} 1_{[l_j, r_j)}(y)\,\frac{\hat{p}_j}{w_j}, \quad 0 \le y < 1, \qquad (73.13)$$

where J is the number of histogram cells, $[l_j, r_j)$ is the support of cell j with width $w_j = r_j - l_j > 0$, and $\hat{p}_j$ is the estimator of the probability of cell j. The corresponding distribution function, $\hat{H}_Y$, is a piecewise linear function, obtained by integrating Equation (73.13) with respect to y, and its inverse, $\hat{H}_Y^{-1}$, is the piecewise linear function

$$\hat{H}_Y^{-1}(x) = \sum_{j=1}^{J} 1_{[\hat{C}_{j-1}, \hat{C}_j)}(x)\left[l_j + \left(x - \hat{C}_{j-1}\right)\frac{w_j}{\hat{p}_j}\right], \quad 0 \le x < 1, \qquad (73.14)$$

where $\{\hat{C}_j\}_{j=0}^{J}$ is the cdf of $\{\hat{p}_j\}_{j=1}^{J}$, i.e., $\hat{C}_j = \sum_{i=1}^{j} \hat{p}_i$, $1 \le j \le J$ $(\hat{C}_0 = 0,\ \hat{C}_J = 1)$.

73.2.6 Stitching Transformations

It turns out that a distortion of the form $D = \hat{H}_Y^{-1}$ is rarely adequate in practice, because background ARM processes, being random walks on the unit circle, produce visual "discontinuities" when the background process crosses the circle origin. Furthermore, a monotone transformation (including $\hat{H}_Y^{-1}$) preserves this effect (see Jagerman and Melamed 1992a; Melamed 1993 for a detailed discussion). To circumvent this problem, we use in practice, a 2-stage composite distortion of the form

$$D(x) = \hat{H}_Y^{-1}(S_\xi(x)), \qquad (73.15)$$

where the first stage, S_ξ, "smoothes" the sample paths of ARM processes (with the uniform marginal distributions preserved intact), while the second stage, $\hat{H}_Y^{-1}$, carries out inversion. The first stage is called a *stitching transformation* and is given by

$$S_\xi(u) = \begin{cases} u/\xi, & 0 \le u \le \xi \\ (1-u)/(1-\xi), & \xi \le u \le 1 \end{cases} \qquad (73.16)$$

where the parameter $\xi \in [0, 1]$ is called a *stitching parameter*. We point out that the smoothing action of stitching transformations is effected for $0 < \xi < 1$, due to their continuity on the unit circle (the term "stitching" is motivated by the fact that $S_\xi(0) = S_\xi(1-)$ for all $0 < \xi < 1$). Note that for the extremal cases, $\xi = 1$ and $\xi = 0$, the corresponding stitching transformations are $S_1(u) = u$ (the identity), and $S_0(u) = 1 - u$ (the antithetic), and these are "non-smoothing," as they are not continuous on the unit circle. Practical modeling most often employs $0 < \xi < 1$, and primarily, the midpoint $\xi = 0.5$. For a discussion of the smoothing properties of stitching transformations and their uniformity-preserving property, refer to Melamed (1991), Jagerman and Melamed (1992a) and Melamed (1993). A geometric interpretation of the stitching parameter is more involved, but can be readily seen when the background process is highly autocorrelated and cyclical (has a drift about the unit circle). In this case, the stitching transformation curve can be seen to serve as the "profile" of a cycle, in the sense that the sample paths of the background process tend to "hug" that curve. The ξ parameter then specifies the apex of the cycle profile.

The introduction of stitching as a smoothing operation still allows us to calculate the Laplace transform of the resulting composite distortion, needed in Equations (73.11–73.12). More specifically, for any general composite distortion of the form

$$D_\xi(u) = D(S_\xi(u)) \qquad (73.17)$$

where D is any distortion, the corresponding Laplace transform is

$$\tilde{D}_\xi(i\,2\pi\nu) = \xi\,\tilde{D}(i\,2\pi\nu\xi) + (1-\xi)\,\tilde{D}(-i\,2\pi\nu(1-\xi)), \quad \nu \ge 0 \qquad (73.18)$$

(Jagerman and Melamed 1992a). Thus, if the original distortion $\tilde{D}$ is known, then any composite distortion, $\tilde{D}_\xi$, is readily computable as well. These have been computed in Jagerman and Melamed (1992b) for various distributions, including the histogram distortion $D = \hat{H}_Y^{-1}$ in Equation (73.14).

73.3 The ARM Modeling Methodology

Suppose the modeler wishes to fit an ARM model to some empirical record $\{Y_n\}_{n=0}^{N}$, from which the empirical distribution (histogram) and empirical autocorrelation function have been estimated. The ARM modeling methodology searches for an innovation sequence, $V = \{V_n\}$, and a stitching parameter, ξ, such that the corresponding ARM process gives rise to an autocorrelation function that approximates well its empirical counterpart; recall that ARM theory ensures that the associated marginal distribution is automatically matched to its empirical counterpart. Such ARM modeling may be implemented as a heuristic interactive search with extensive modeler interaction, or as an algorithmic search with minimal modeler intervention. Both approaches require software support (Hill and Melamed 1995). The ARM modeling methodology unfolds in practice as follows:

- First, the ARM process sign (ARM^+ or ARM^-) is selected. Such a selection is typically guided by the modeler's experience (however, trying both classes would simply double the modeling time).
- Next, the empirical inverse distribution function, $\hat{H}_Y^{-1}$, is constructed from the empirical data. Once the modeler specifies the number of histogram cells and their width, the distortion transformation is readily computed from Equation (73.14) and Equation (73.18).
- The third step, and the hardest one, is to search for a pair, (V, ξ), consisting of an innovation process and a stitching parameter, which jointly give rise to an ARM process that approximates well the leading empirical autocorrelations.

In the heuristic approach, the modeler embarks on a heuristic search of pairs (V, ξ), guided by visual feedback (Melamed 1993; Hill and Melamed 1995). This approach, however, may be time-consuming and consequently can engender considerable tedium. Alternatively, an algorithmic search approach can be devised, analogously to Jelenkovic and Melamed (1995). In this approach, a search is combined with nonlinear optimization to produce a set of ARM candidate models. First, the parameter space, (V, ξ), is quantized into a moderately large, but finite, set Q of ARM model specifications. The autocorrelation function is computed for each model in Q via the computationally fast formulas in Equations (73.11) and (73.12). A "best set" of m models is selected from Q (the value of m would be typically small, on the order of 10); the associated "goodness-of-fit" criterion would select those ARM models whose autocorrelation functions have the smallest distance from their empirical counterpart in some prescribed metric. A standard choice of such a metric is the weighted sum of the squared autocorrelation differences, the sum ranging over some span of lags, $\tau = 1, \ldots, L$. Next, a Steepest-Descent nonlinear optimization is performed for models in the "best set" only, to further reduce the aforementioned autocorrelation distance. The modeler then selects the "best of the best," either as the model with the overall smallest distance, or an optimized model in the "best set" that gives rise to Monte Carlo sample paths, which appear most "similar" to the empirical data.

73.4 The ARM Forecasting Methodology

The computational basis of the ARM forecasting methodology is provided by the transition densities in Equations (73.9–73.10). These equations allow us to compute numerically conditional densities $f_{U^+_{k+\tau}|U^+_k}(v|u)$ and $f_{U^+_{k+\tau}|U^+_k}(v|u)$ of ARM background processes, and consequently, their foreground counterparts, $f_{X^+_{k+\tau}|U^+_k}(x|u)$ and $f_{X^-_{k+\tau}|U^+_k}(x|u)$, can be readily computed too, in view of Equations (73.7–73.8), which relate background and foreground ARM processes via a deterministic transformation. However, we cannot readily compute a transition density of foreground ARM processes of the form $f_{X^+_{k+\tau}|X^+_k}(x|y)$ and $f_{X^-_{k+\tau}|X^-_k}(x|y)$. The problem stems from the fact that for $0 < \xi < 1$ and a given stitched foreground ARM value, x, the stitching transformation (Equation 73.16) has two background ARM value solutions,

$$u^{(1)}(x) = \xi\, x,\; u^{(2)}(x) = 1 - (1-\xi)\, x. \tag{73.19}$$

To get around this problem, consider any background ARM process $\{U_n\}_{n=0}^{\infty}$ and its foreground counterpart, $\{X_n\}_{n=0}^{\infty}$. ARM forecasting uses systematically a convex (probabilistic) combination of transition densities of the form

$$\begin{aligned} c_{k\pm\tau|k}(y|x) &= p\, f_{X_{k\pm\tau}|U_k}(y|u^{(1)}(x)) \\ &\quad + (1-p)\, f_{X_{k\pm\tau}|U_k}(y|u^{(2)}(x)) \end{aligned} \tag{73.20}$$

for some probability value $0 \le p \le 1$, referred to as the *mixing parameter*, whose selection via a simple optimization procedure will be described later. More specifically, using the transition density of Equation (73.20), ARM forecasts consist of both point estimates and confidence intervals as follows:

- A point estimate of the future value of $X_{k+\tau}$ given the current value of X_k is computed as the conditional expectation with respect to the density $c_{k+\tau|k}(y|x)$

$$\begin{aligned} &\mathrm{E}_{c_{k+\tau|k}}[X_{k+\tau} \mid X_k = x] \\ &\quad = p\,\mathrm{E}\left[X_{k+\tau}|U_k = u^{(1)}(x)\right] \\ &\qquad + (1-p)\,\mathrm{E}\left[X_{k+\tau}|U_k = u^{(2)}(x)\right] \end{aligned} \tag{73.21}$$

- A confidence interval with a prescribed probability about the point estimate (Equation (73.21)) is computed via an appropriate transition density $c_{k+\tau|k}(y|x)$ of the form of Equation (73.20).

We next proceed to describe the details of the ARM forecasting methodology.

73.4.1 Selection of the Mixing Parameter

The selection of the mixing parameter utilizes time reversal (Kelly 1979) of the background ARM process, thereby generalizing the approach of Jagerman and Melamed (1995). This approach calls for the computation of a prescribed number of *backward* (time reversed) conditional expectations of the form

$$\begin{aligned} &\mathrm{E}_{c_{k-\tau|k}}\left[X_{k-\tau} \mid X_k = x\right] \\ &\quad = p\,\mathrm{E}\left[X_{k-\tau}|U_k = u^{(1)}(x)\right] \\ &\qquad + (1-p)\,\mathrm{E}\left[X_{k-\tau}|U_k = u^{(2)}(x)\right] \end{aligned} \tag{73.22}$$

To this end, we take advantage of the fact that the Markovian property of background ARM processes is preserved under time reversal, and the backward transition densities are readily given in terms of the forward transition densities. Accordingly, for any ARM background process $\{U_n\}_{n=0}^{\infty}$ and $0 \le u, v < 1,\ \tau \ge 1$,

$$f_{U_{k-\tau}|U_k}(v|u) = f_{U_k|U_{k-\tau}}(u|v)\,\frac{f_{U_{k-\tau}}(v)}{f_{U_k}(u)} = f_{U_k|U_{k-\tau}}(u|v) \tag{73.23}$$

because the marginal density of ARM background processes is uniform on [0,1). Using Equations (73.9–73.10), we can write the formulas for backward transition densities as follows. For any background ARM$^+$ process $\{U_n^+\}$ and $0 \le u, v < 1, \tau \ge 1$,

$$\begin{aligned} f_{U_{k-\tau}^+|U_k^+}(v|u) &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,\,k}}(i\,2\,\pi\,\nu)\, e^{i\,2\,\pi\,\nu\,(u-v)} \\ &= 1 + 2\sum_{\nu=1}^{\infty} \mathrm{Re}\left[\tilde{f}_{S_{k-\tau+1,\,k}}(i\,2\,\pi\,\nu)\, e^{i\,2\,\pi\,\nu\,(u-v)}\right] \end{aligned} \tag{73.24}$$

and for any background ARM$^-$ process $\{U_n^-\}$ and $0 \le u, v < 1, \tau \ge 1$,

$$f_{U_{k-\tau}^-|U_k^-}(v|u) = \begin{cases} \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,\,k}}(i\,2\,\pi\,\nu)\, e^{i\,2\,\pi\,\nu\,(u-v)}, & k \text{ even},\ \tau \text{ even} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,\,k}}(i\,2\,\pi\,\nu)\, e^{i\,2\,\pi\,\nu\,(u+v)}, & k \text{ even},\ \tau \text{ odd} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,\,k}}(i\,2\,\pi\,\nu)\, e^{i\,2\,\pi\,\nu\,(-u+v)}, & k \text{ odd},\ \tau \text{ even} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,\,k}}(i\,2\,\pi\,\nu)\, e^{i\,2\,\pi\,\nu\,(-u-v)}, & k \text{ odd},\ \tau \text{ odd} \end{cases} \tag{73.25}$$

Equation (73.25) can be simplified analogously to Equation (73.24).

Observing that past values are known in practice, the idea is to generate retrograde point estimates and compare them to a window of (known) past values. The mixing parameter p is then selected so as to minimize an objective function involving deviations of retrograde forecasts from known past values, analogously to ordinary regression. More specifically, for each time index k, consider the objective function

$$g_k(p) = \sum_{\tau=1}^{N} w_\tau \left[p e_{k-\tau}^{(1)}(\hat{x}_k) + (1-p)\, e_{k-\tau}^{(2)}(\hat{x}_k) - \hat{x}_{k-\tau}\right]^2 \tag{73.26}$$

where p is the mixing parameter to be selected, N is the window size, the w_τ are given non-negative weights, $e_k^{(i)}(\hat{x}_k) = \mathbf{E}\left[D(U_{k-\tau})|U_k = u^{(i)}(\hat{x}_k)\right]$, for $i = 1, 2$, and $\hat{x}_k$ is the observed value of X_k. Our goal is to select a mixing parameter p that minimizes Equation (73.26) by differentiating it with respect to p, setting the derivative to zero and then solving for p. The nominal solution is given by

$$p = \frac{\sum\limits_{\tau=1}^{N} w_\tau \left[e_{k-\tau}^{(1)}(\hat{x}_k) - e_{k-\tau}^{(2)}(\hat{x}_k)\right]\left[\hat{x}_{k-\tau} - e_{k-\tau}^{(2)}(\hat{x}_k)\right]}{\sum\limits_{\tau=1}^{N} w_\tau \left[e_{k-\tau}^{(1)}(\hat{x}_k) - e_{k-\tau}^{(2)}(\hat{x}_k)\right]^2} \tag{73.27}$$

If Equation (73.27) yields a value between 0 and 1, then p is selected as the actual mixing parameter (if the denominator vanishes, select any value of p, say, $p = 1$). Otherwise, if it yields a negative value, then the actual mixing parameter is set to $p = 0$, and if it yields a value that exceeds 1, it is set to $p = 1$.

73.4.2 Computation of Conditional Expectations

Conditional expectations are used in the computation of the point forecasts in Equations (73.21) and (73.22). We begin with the computation of forward conditional expectations, utilizing Equation (73.9). For any background ARM$^+$ process $\{U_n^+\}$ and $\tau \geq 1$,

$$
\begin{aligned}
\mathrm{E}\left[X_{k+\tau}^{+} | U_k^{+} = u\right] &= \int_0^1 D(\nu)\, f_{U_{k+\tau}^{+}|U_k^{+}}(\nu|u)\, d\nu \\
&= \int_0^1 D(\nu) \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\, e^{i2\pi\nu(\nu-u)}\, d\nu \\
&= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu) \int_0^1 D(\nu)\, e^{i2\pi\nu(\nu-u)}\, d\nu \\
&= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\, e^{-i\,2\,\pi\,\nu\,u}\, \tilde{D}(-i\,2\,\pi\,\nu) \\
&= \mu_X + 2\sum_{\nu=1}^{\infty} \mathrm{Re}\left[\tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\, e^{-i2\pi\nu u}\, \tilde{D}(-i\,2\,\pi\,\nu)\right]
\end{aligned}
\tag{73.28}
$$

Utilizing Equation (73.10), a similar computation for any foreground ARM$^-$ process, $\{X_n^-\}$, and $\tau \geq 1$ yields

$$
E\left[X_{k+\tau}^{-} | U_k^{-} = u\right] = \begin{cases}
\sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\, e^{-i2\pi\nu u}\, \tilde{D}(-i\,2\,\pi\,\nu), & k \text{ even},\ \tau \text{ even} \\
\sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\, e^{-i2\pi\nu u}\, \tilde{D}(i\,2\,\pi\,\nu), & k \text{ even},\ \tau \text{ odd} \\
\sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\, e^{i2\pi\nu u}\, \tilde{D}(i\,2\,\pi\,\nu), & k \text{ odd},\ \tau \text{ even} \\
\sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,k+\tau}}(i\,2\,\pi\,\nu)\, e^{i2\pi\nu u}\, \tilde{D}(-i\,2\,\pi\,\nu), & k \text{ odd},\ \tau \text{ odd}
\end{cases}
\tag{73.29}
$$

Equation (73.29) can be further simplified analogously to the last line in Equation (73.28).

The computation of backward conditional expectations utilizes Equations (73.24) and (73.25), and is similar to that in Equation (73.28). For any background ARM$^+$ process $\{U_n^+\}$ and $\tau \geq 1$,

$$
\begin{aligned}
E\left[X_{k-\tau}^{+} | U_k^{+} = u\right] &= \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\,\pi\,\nu)\, e^{i2\pi\nu u}\, \tilde{D}(i\,2\,\pi\,\nu) \\
&= \mu_X + 2\sum_{\nu=1}^{\infty} \mathrm{Re}\left[\tilde{f}_{S_{k-\tau+1,k}}(i\,2\,\pi\,\nu)\, e^{i2\pi\nu u}\, \tilde{D}(i\,2\,\pi\,\nu)\right]
\end{aligned}
\tag{73.30}
$$

and for any foreground ARM$^-$ process, $\{X_n^-\}$, and $\tau \geq 1$,

$$E\left[X_{k-\tau}^- | U_k^- = u\right] = \begin{cases} \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\pi\,\nu)\, e^{i\,2\pi\,\nu\,u}\, \tilde{D}(i\,2\pi\,\nu), & k \text{ even}, \ \tau \text{ even} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\pi\,\nu)\, e^{i\,2\pi\,\nu\,u}\, \tilde{D}(-i\,2\pi\,\nu), & k \text{ even}, \ \tau \text{ odd} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\pi\,\nu)\, e^{-i\,2\pi\,\nu\,u}\, \tilde{D}(-i\,2\pi\,\nu), & k \text{ odd}, \ \tau \text{ even} \\ \sum\limits_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\pi\,\nu)\, e^{-i\,2\pi\,\nu\,u}\, \tilde{D}(i\,2\pi\,\nu), & k \text{ odd}, \ \tau \text{ odd} \end{cases} \tag{73.31}$$

Equation (73.31) can be further simplified analogously to Equation (73.30).

73.4.3 Computation of Conditional Distributions

Conditional cumulative distribution functions (cdf) are used in the computation of confidence intervals for forecasts. For any foreground ARM process, $\{X_n\}_{n=0}^{\infty}$, we note the relations

$$F_{X_{k\pm\tau}|U_k}(x|u) = F_{S_\xi(U_{k\pm\tau})|U_k}(F_X(x)|u) \tag{73.32}$$

where $S_\xi(U_{k\pm\tau})$ is a stitched background random variable and F_X is the common cdf of the foreground ARM random variables. Equation (73.32) shows that it suffices to derive the formulas for the conditional cdf $F_{S_\xi(U_{k\pm\tau})|U_k}(s|u)$, as the requisite one is readily obtained by setting $s = F_X(x)$ in Equation (73.32). Accordingly, utilizing Equation (73.9), one has for any background ARM$^+$ process $\{U_n^+\}$ and $\tau \geq 1$,

$$\begin{aligned} F_{S_\xi(U_{k+\tau}^+)|U_k^+}(s|u) &= \int_0^s f_{S_\xi(U_{k+\tau}^+)|U_k^+}(r|u)\,dr \\ &= 1 - \int_{u^{(1)}(s)}^{u^{(2)}(s)} f_{U_{k+\tau}^+|U_k^+}(v|u)\,dv \\ &= 1 - \int_{u^{(1)}(s)}^{u^{(2)}(s)} \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\pi\,\nu)\, e^{i\,2\pi\,\nu\,(v-u)}\,dv \\ &= 1 - \sum_{\nu=-\infty}^{\infty} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\pi\,\nu)\, e^{-i\,2\pi\,\nu\,u} \int_{u^{(1)}(s)}^{u^{(2)}(s)} e^{i\,2\pi\,\nu\,v}\,dv \\ &= s + \sum_{\nu\neq 0} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\pi\,\nu)\, \frac{e^{i\,2\pi\,\nu\,(-u+u^{(1)}(s))} - e^{i\,2\pi\,\nu\,(-u+u^{(2)}(s))}}{i\,2\pi\,\nu} \\ &= s + 2\sum_{\nu=1}^{\infty} \mathrm{Re}\left[\tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\pi\,\nu)\, \frac{e^{i\,2\pi\,\nu\,(-u+u^{(1)}(s))} - e^{i\,2\pi\,\nu\,(-u+u^{(2)}(s))}}{i\,2\pi\,\nu}\right] \end{aligned} \tag{73.33}$$

Utilizing Equation (73.10), a similar computation for any foreground ARM$^-$ process, $\{X_n^-\}$, and $\tau \geq 1$ yields

$$F_{S_\xi(U_{k+\tau}^-)|U_k^-}(s|u) = \int_0^s f_{S_\xi(U_{k+\tau}^-)|U_k^-}(r|u)\,dr$$
$$= \begin{cases} s + \sum_{\nu \neq 0} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(-u+u^{(1)}(s))} - e^{i\,2\,\pi\,\nu\,(-u+u^{(2)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ even, } \tau \text{ even} \\ s + \sum_{\nu \neq 0} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(-u-u^{(2)}(s))} - e^{i\,2\,\pi\,\nu\,(-u-u^{(1)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ even, } \tau \text{ odd} \\ s + \sum_{\nu \neq 0} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(u-u^{(2)}(s))} - e^{i\,2\,\pi\,\nu\,(u-u^{(1)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ odd, } \tau \text{ even} \\ s + \sum_{\nu \neq 0} \tilde{f}_{S_{k+1,\,k+\tau}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(u+u^{(1)}(s))} - e^{i\,2\,\pi\,\nu\,(u+u^{(2)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ odd, } \tau \text{ odd} \end{cases} \tag{73.34}$$

Equation (73.34) can be further simplified analogously to the last line in Equation (73.33).

The computation of backward conditional distributions utilizes Equations (73.24) and (73.25), and is similar to that in Equations (73.33) and (73.34). For any background ARM$^+$ process $\{U_n^+\}$ and $\tau \geq 1$,

$$F_{S_\xi(U_{k-\tau}^+)|U_k^+}(s|u) = \int_0^s f_{S_\xi(U_{k-\tau}^+)|U_k^+}(r|u)\,dr$$
$$= s + \sum_{\nu \neq 0} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\,\pi\,\nu)\,\frac{e^{i\,2\,\pi\,\nu\,(u-u^{(2)}(s))} - e^{i\,2\,\pi\,\nu\,(u-u^{(1)}(s))}}{i\,2\,\pi\,\nu} \tag{73.35}$$

and for any foreground ARM$^-$ process, $\{X_n^-\}$, and $\tau \geq 1$,

$$F_{S_\xi(U_{k-\tau}^-)|U_k^-}(s|u) = \int_0^s f_{S_\xi(U_{k-\tau}^-)|U_k^-}(r|u)\,dr$$
$$= \begin{cases} s + \sum_{\nu \neq 0} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(u-u^{(2)}(s))} - e^{i\,2\,\pi\,\nu\,(u-u^{(1)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ even, } \tau \text{ even} \\ s + \sum_{\nu \neq 0} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(u+u^{(1)}(s))} - e^{i\,2\,\pi\,\nu\,(u+u^{(2)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ even, } \tau \text{ odd} \\ s + \sum_{\nu \neq 0} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(-u+u^{(1)}(s))} - e^{i\,2\,\pi\,\nu\,(-u+u^{(2)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ odd, } \tau \text{ even} \\ s + \sum_{\nu \neq 0} \tilde{f}_{S_{k-\tau+1,k}}(i\,2\,\pi\,\nu)\,\dfrac{e^{i\,2\,\pi\,\nu\,(-u-u^{(2)}(s))} - e^{i\,2\,\pi\,\nu\,(-u-u^{(1)}(s))}}{i\,2\,\pi\,\nu}, & k \text{ odd, } \tau \text{ odd} \end{cases} \tag{73.36}$$

Again, Equations (73.35) and (73.36) can be further simplified analogously to the last line in Equation (73.33).

We conclude this section by observing that the conditional densities of $F_{S_\xi(U_{k\pm\tau})|U_k}(s|u)$ in Equations (73.33–73.36) are readily obtained by differentiating them with respect to s. The transition densities for any foreground ARM process $\{X_n\}_{n=0}^{\infty}$ are then given by the formula

$$f_{X_{k\pm\tau}|U_k}(x|u) \;=\; f_{S_\xi(U_{k\pm\tau})|U_k}(F_X(x)|u)\, f_X(x) \qquad (73.37)$$

Finally, the conditional densities $f_{S_\xi(U_{k\pm\tau})|U_k}(s|u)$ can be readily expressed in terms of the conditional densities $f_{U_{k\pm\tau}|U_k}(v|u)$. To this end, write

$$\begin{aligned} F_{S_\xi(U_{k\pm\tau})|U_k}(s|u) &= F_{U_{k\pm\tau}|U_k}(u^{(1)}(s)|u) \\ &\quad +1 - F_{U_{k\pm\tau}|U_k}(u^{(2)}(s)|u) \\ &= F_{U_{k\pm\tau}|U_k}(s\xi|u) \\ &\quad +1 - F_{U_{k\pm\tau}|U_k}(1-(1-\xi)s|u) \end{aligned}$$

and differentiating this equation with respect to s yields

$$\begin{aligned} f_{S_\xi(U_{k\pm\tau})|U_k}(s|u) &= \xi\, f_{U_{k\pm\tau}|U_k}(s\xi|u) \\ &\quad + (1-\xi)\, f_{U_{k\pm\tau}|U_k}(1-(1-\xi)s|u) \end{aligned} \qquad (73.38)$$

73.5 Example: ARM Modeling of an S&P 500 Time Series

Stochastic financial models are used to support investment decisions in financial instruments: short-term forecasting is employed to guide day-to-day trading, while longer-term planning (e.g., comparing alternative investment strategies) often calls for the generation of financial Monte Carlo scenarios. Furthermore, quantitative investment strategies require scenarios to estimate internal parameters of various financial models (Hull 2009), typically to valuate the current fair price of a security, rather than to forecast its future price. As a time series modeling and forecasting methodology, ARM processes are suited for both purposes.

The modeling methodology described in Sect. 73.3 was applied to an empirical time series of 500 values drawn from the S&P 500 Index, yielding an ARM$^+$ model. Figure 73.1 displays the best model found by the ARM modeling methodology.

The screen image in Fig. 73.1 consists of four panels with various statistics providing visual information on the fit quality as follows:

1. The upper-left panel depicts the empirical S&P 500 Index time series and a simulated time series of the fitted model. The initial value of both time series were arranged to coincide.
2. The upper-right panel depicts an empirical histogram of the S&P 500 Index time series with a compatible histogram of the simulated time series superimposed on it.
3. The lower-left panel depicts the empirical autocorrelation function of the S&P 500 Index data, as well as its simulated and theoretical counterparts.
4. The lower-right panel depicts the empirical spectral density function of the S&P 500 Index data, as well as its simulated and theoretical counterparts.

An inspection of Fig. 73.1 reveals that the fitted ARM$^+$ model provides a good fit to the empirical data in accordance with requirements (a)–(c) in Sect. 73.1.

Next, the fitted ARM$^+$ model was exercised to compute point forecasts (conditional expectations), using the ARM forecasting methodology as described in Sects. 73.4.1 and 73.4.2. Figure 73.2 displays various forecasting-related statistics for the empirical sample of S&P Index data.

The screen image in Fig. 73.2 consists of four panels with various statistics pertaining to point forecasts as follows:

1. The upper-left panel depicts the tail-end of the empirical S&P 500 Index time series with forward forecasts superimposed on it. Each empirical value “sprouts” a branch consisting of a sequence of 3 forward forecasts, computed by optimizing the backward fitting of the previous 5 empirical values (if available).
2. The upper-right panel depicts the tail-end of the empirical S&P 500 Index time series and the corresponding sequence of lag-1 forward forecasts. The deviation of a forecast value from the (empirical) true value can be gauged by the vertical distances along the two time series.
3. The lower-left panel depicts the sequence of the aforementioned deviation values.
4. The lower-right panel depicts a histogram of the aforementioned deviation values.

An inspection of Fig. 73.2 shows that the forecasts are quite good. In particular, inspection of the last panel shows that the deviations are arranged in a tight range around the origin, supporting this observation.

Figure 73.3 depicts a magnified view of the upper-left panel in Fig. 73.2.

Note that the forecasting branches lie quite closely to the empirical data, indicating good forecasting results.

Figure 73.4 depicts a magnified view of the upper-right panel in Fig. 73.2.

Again the two curves of the (empirical data and forecasts) lie quite close to each other. A single out-of-sample forecast is shown in the upper right corner (unmatched by empirical data).

Finally, the fitted ARM$^+$ model was exercised to compute forecast distributions (conditional distributions and

Fig. 73.1 Statistics of the best model fitted to a sample of S&P 500 Index data

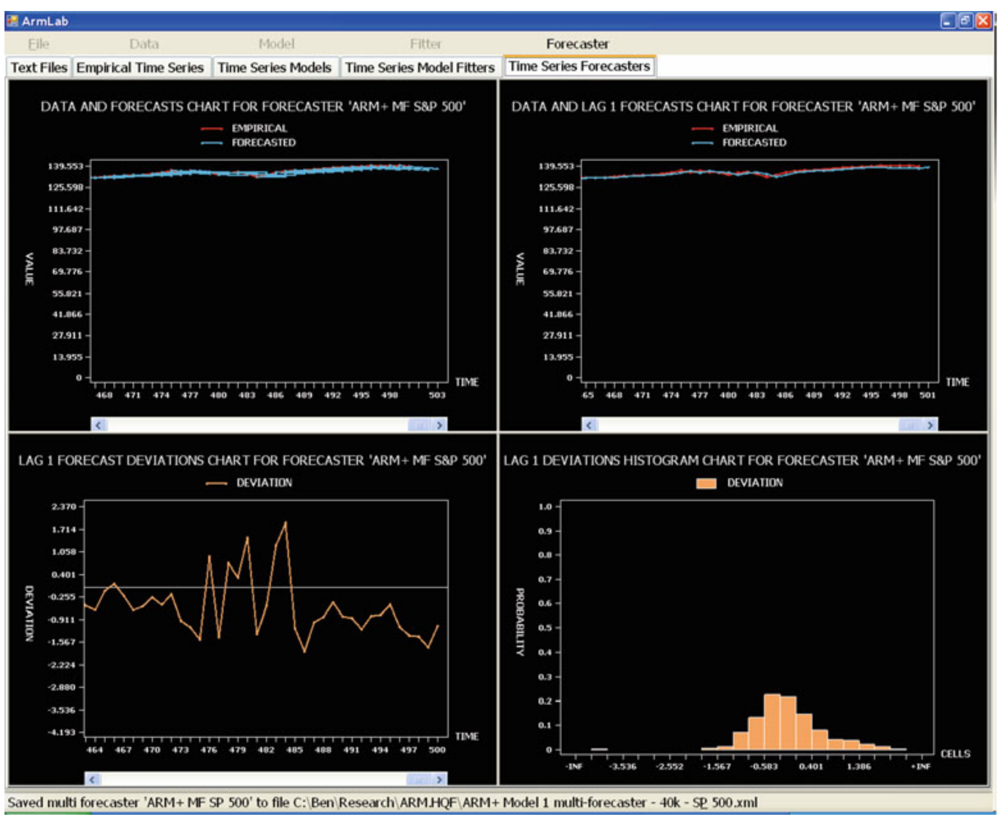


Fig. 73.2 Multiple-Lag forecasts of the best model fitted to a sample of S&P 500 Index data

Fig. 73.3 Multiple-lag forecasts for a sample of S&P 500 Index data

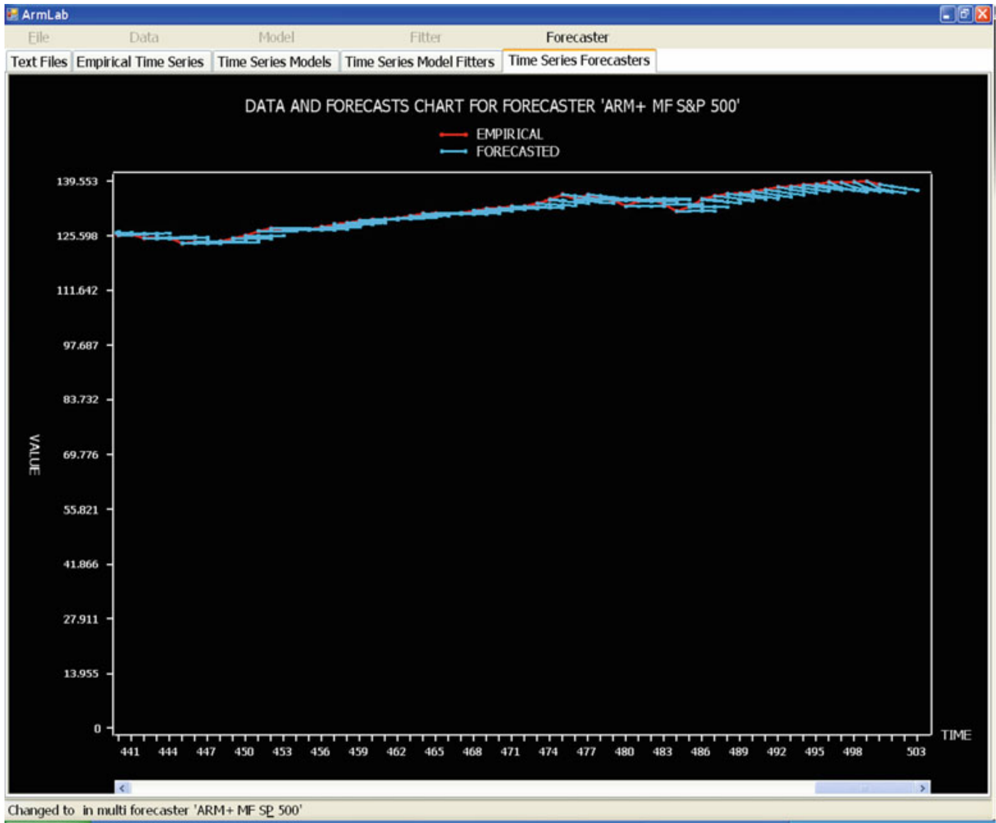
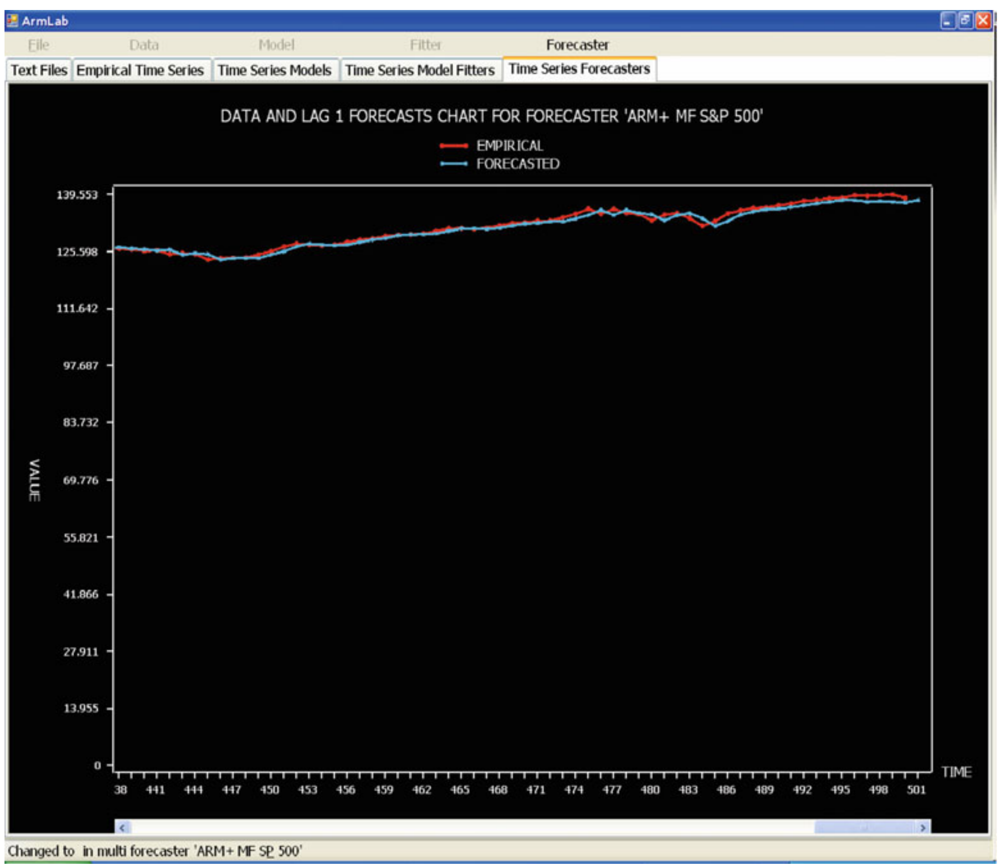


Fig. 73.4 Single-lag forecasts for a sample of S&P 500 Index data



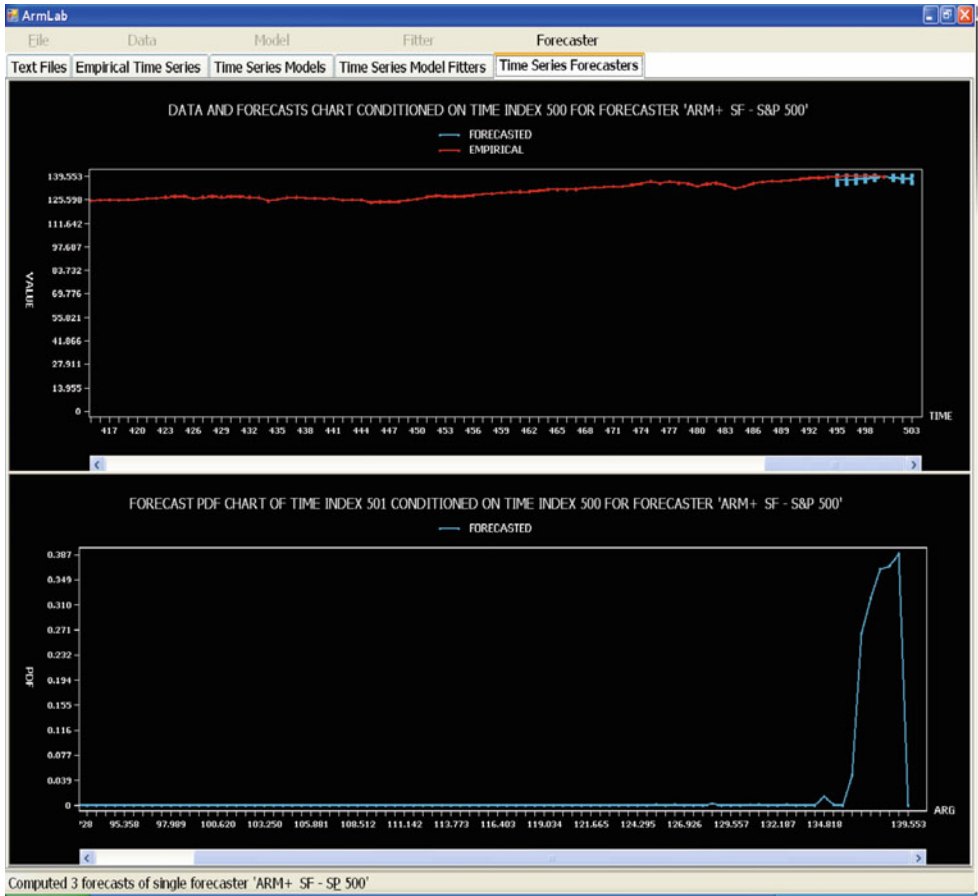


Fig. 73.5 Forecasts of the best model fitted to a sample of S&P 500 Index data for the last value of the empirical data

densities), as described in Sect. 73.4.3. Figure 73.5 depicts these statistics for the last empirical sample value of the S&P 500 Index data.

The screen image in Fig. 73.5 displays two panels with forecast information as follows:

1. The upper panel depicts the empirical S&P 500 Index time series with forward and backward forecasts superimposed on it for the last empirical value. Shown are the sequence of 3 forward forecasts and 5 backward forecasts, as well as the 90% confidence interval about each of them.
2. The lower panel depicts the density of the corresponding forward lag-1 forecast, conditioned on the last empirical value. Accordingly, the confidence interval for the lag-1 forward forecast was obtained by a simple search over the values of the cdf corresponding to the displayed pdf.

Recall that the forward forecast values in Fig. 73.5 constitute out-of-sample predictions on future values of the S&P 500 Index.

73.6 Summary

The class of ARM (Autoregressive Modular) processes is a class of stochastic processes, defined by a nonlinear autoregressive scheme with modulo-1 reduction and additional transformations. ARM processes constitute a versatile class designed to produce high-fidelity models from stationary empirical time series by fitting a strong statistical signature consisting of the empirical marginal distribution (histogram) and the empirical autocorrelation function. More specifically, fitted ARM processes guarantee the matching of arbitrary empirical distributions, and simultaneously permit the approximation of the leading empirical autocorrelations. Additionally, simulated sample paths of ARM models often resemble the data to which they were fitted. Thus, ARM processes aim to provide both a quantitative and qualitative fit to empirical data. Fitted ARM models can be used in two ways: (1) to generate realistic-looking Monte Carlo sample paths (e.g., financial scenarios), and (2) to forecast via point estimates as well as confidence intervals.

This chapter starts with a high-level discussion of stochastic model fitting and explains the fitting approach that motivates the introduction of ARM processes. It then continues with a review of ARM processes and their fundamental properties, including their construction, transition structure, and autocorrelation structure. Next, the chapter proceeds to outline the ARM modeling methodology by describing the key steps in fitting an ARM model to empirical data. It then describes in some detail the ARM forecasting methodology, and the computation of the conditional expectations that serve as point estimates (forecasts of future values) and their underlying conditional distributions from which confidence intervals are constructed for the point estimates. Finally, the chapter concludes with an illustration of the efficacy of the ARM modeling and forecasting methodologies utilizing a sample from the S&P 500 Index.

Acknowledgements I would like to thank Dr. David Jagerman, Xiang Zhao, and Junmin Shi for reading and commenting on the manuscript.

References

Altiok, T. and B. Melamed. 2001. “The case for modeling correlation in manufacturing systems.” *IIE Transactions* 33(9), 779–791.

Bratley, P., B. L. Fox, and L. E. Schrage. 1987. *A guide to simulation*, Springer, New York.

Cario, M. C. and B. L. Nelson. 1996. “Autoregressive to anything: time-series input processes for simulation.” *OR Letters* 19, 51–58.

Diaconis, P. 1988. *Group representations in probability and statistics*, Lecture Notes – Monograph Series 11, Institute of Mathematical Statistics, Hayward, CA.

Hill, J. R. and B. Melamed. 1995. “TEStool: a visual interactive environment for modeling autocorrelated time series.” *Performance Evaluation* 24(1&2), 3–22.

Hull, J. 2009. *Options, futures, and other derivative securities*, 7th ed., Prentice-Hall, Upper Saddle River.

Jagerman, D. L. and B. Melamed. 1992a. “The transition and autocorrelation structure of TES processes part I: general theory.” *Stochastic Models* 8(2), 193–219.

Jagerman, D. L. and B. Melamed. 1992b. “The transition and autocorrelation structure of TES processes part II: special cases.” *Stochastic Models* 8(3), 499–527.

Jagerman, D. L. and B. Melamed. 1994. “The spectral structure of TES processes.” *Stochastic Models* 10(3), 599–618.

Jagerman, D. L. and B. Melamed. 1995. “Bidirectional estimation and confidence regions for TES processes.” *Proceedings of MASCOTS ‘95*, Durham, NC, pp. 94–98.

Jelenkovic, P. and B. Melamed. 1995. “Algorithmic modeling of TES processes.” *IEEE Transactions on Automatic Control* 40(7), 1305–1312.

Kelly, F. P. 1979. *Reversibility and stochastic networks*, Wiley, New York.

Klaoudatos, G., M. Devetsikiotis, and I. Lambadaris. 1999. *“Automated modeling of broadband network data using the QTES methodology,” Proceedings IEEE ICC ‘99*, Vol 1, Vancouver, Canada, pp. 397–403.

Law, A. and D. W. Kelton. 1991. *Simulation modeling and analysis*, McGraw-Hill, New York.

Livny, M., B. Melamed, and A. K. Tsiolis. 1993. “The impact of autocorrelation on queueing systems.” *Management Science* 39(3), 322–339.

Melamed, B. 1991. “TES: a class of methods for generating autocorrelated uniform variates.” *ORSA Journal on Computing* 3(4), 317–329.

Melamed, B. 1993. “An overview of TES processes and modeling methodology,” in *Performance evaluation of computer and communications systems*, L. Donatiello and R. Nelson (Eds.). Lecture Notes in Computer Science, Springer, New York, pp. 359–393.

Melamed, B. 1997. “The empirical TES methodology: modeling empirical time series.” *Journal of Applied Mathematics and Stochastic Analysis* 10(4), 333–353.

Melamed, B. 1999. “ARM processes and modeling methodology.” *Stochastic Models* 15(5), 903–929.

Melamed, B., Q. Ren, and B. Sengupta. 1996. “The QTES/PH/1 queue,” *Performance Evaluation* 26, 1–20.