Chapter 3
Forecasting, Confidence Band Estimation and Updating

Abstract In this chapter, forecasting, forecast confidence band estimation, and the forecast updating methodologies, provided for ARIMA models in the literature, are modified and presented for the ARFIMA models.

Keywords Forecasting, • Confidence band estimation • Updating • ARFIMA models • Confidence limits

3.1 Forecasting

Forecasting and updating of ARFIMA processes are a natural extension of those of ARIMA models. When fractional difference parameter d is 0, Eq. (2.2) represents the ARMA processes and when it is integer, Eq. (2.2) represents the ARIMA processes. Box and Jenkins (1976) provide the details on the forecasting and updating of classic ARIMA models. Forecasting as a conditional expectation of X_{t+l} is said to be made at origin t for a lead time $l \geq 1$ when written as an infinite sum of previous observations plus a random shock;

$$\left[X_{t+l}\right] = \hat{X}_t(l) = \sum_{j=1}^{\infty} \pi_j \left[X_{t+l-j}\right] + \left[\varepsilon_{t+l}\right] \tag{3.1}$$

where π weights may be obtained by equating the coefficients in

$$\phi(B)(1-B)^d = (1 - \pi_1 B - \pi_2 B^2 - \ldots)\theta(B) \tag{3.2}$$

Because of the invertibility condition, the π weights must form a convergent series. Therefore, the forecast is dependent to an important extent only on recent past values (Box and Jenkins 1976). The variance of the forecast error $e_t(l) = X_{t+l} - \hat{X}_t(l)$ is

$$\mathrm{var}(e_t(l)) = \left[1 + \sum_{j=1}^{l-1} \psi_j^2\right] \sigma_\varepsilon^2 \tag{3.3}$$

A. Ercan et al., *Long-Range Dependence and Sea Level Forecasting*, SpringerBriefs in Statistics, DOI: 10.1007/978-3-319-01505-7_3, © The Author(s) 2013

where σ_{ε}^{2} is the variance of the residuals and ψ weights may be obtained by equating coefficients in

$$\phi(B)(1-B)^{d}\left(1+\psi_{1}B+\psi_{2}B^{2}+\ldots\right)=\theta\,(B)\ldots \tag{3.4}$$

Equations. (3.1–3.4), which are valid for ARIMA models (Box and Jenkins 1976), are also valid for the ARFIMA models when π and ψ weights are calculated for the non-integer differencing parameter d from Eqs. (3.2) and (3.4), respectively.

π weights can be obtained by inserting the definitions of the differencing operator $(1-B)^{d}$, as in Eq. (2.3), and the autoregressive and moving-average operators $\phi\,(B)$ and $\theta\,(B)$ into Eq. (3.2);

$$\begin{aligned}\left(1-\phi_{1}B-\phi_{2}B^{2}-\ldots-\phi_{p}B^{p}\right)&\left(1+\frac{\Gamma(1-d)}{\Gamma(-d)\Gamma(2)}B+\frac{\Gamma(2-d)}{\Gamma(-d)\Gamma(3)}B^{2}+\ldots\right)\\&=\left(1-\pi_{1}B-\pi_{2}B^{2}+\ldots\right)\left(1+\theta_{1}B+\theta_{2}B^{2}+\ldots+\theta_{q}B^{q}\right)\end{aligned} \tag{3.5}$$

ψ weights can be obtained similarly by rearranging Eq. (3.4) and using the definitions of differencing and the autoregressive and moving-average operators:

$$\begin{aligned}\left(1+\frac{\Gamma(1+d)}{\Gamma(d)\Gamma(2)}B+\frac{\Gamma(2+d)}{\Gamma(d)\Gamma(3)}B^{2}+\ldots\right)&\left(1+\theta_{1}B+\theta_{2}B^{2}+\ldots+\theta_{q}B^{q}\right)\\&=\left(1-\phi_{1}B-\phi_{2}B^{2}-\ldots-\phi_{p}B^{p}\right)\left(1+\psi_{1}B+\psi_{2}B^{2}+\ldots\right)\end{aligned} \tag{3.6}$$

From Eqs. (3.5–3.6), one can obtain π and ψ weights if fractional difference parameter d is known along with the autoregressive and moving-average coefficients. For example, for ARFIMA(1, d, 0) process when the autoregressive coefficient is ϕ_{1}, π weights can be obtained from the series

$$\pi_{1}=\phi_{1}-\frac{\Gamma(1-d)}{\Gamma(-d)\Gamma(2)}\ \text{and}\ \pi_{j}=\phi_{1}\frac{\Gamma(j-1-d)}{\Gamma(-d)\Gamma(j)}-\frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}\quad \text{for}\,j=2,3\ldots \tag{3.7}$$

and ψ weights can be obtained from the series

$$\psi_{0}=1,\psi_{j}=\phi_{1}\psi_{j-1}+\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\ \text{for}\,j=1,2\ldots \tag{3.8}$$

3.2 Confidence Band Estimation

From the estimation of the variance of the forecast error by Eq. (3.3), one can then estimate the confidence limits Fig. (3.1). Confidence limits can be calculated from

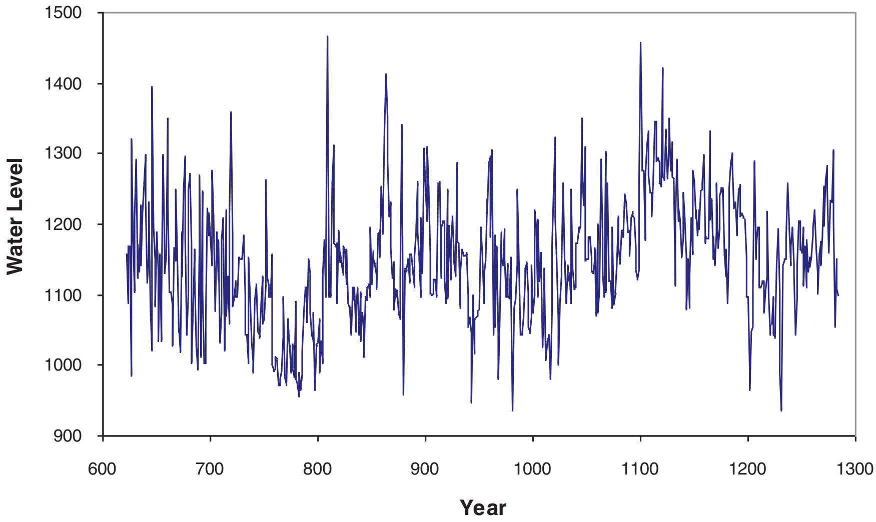

Fig. 3.1 Minimum water levels of Nile River (*source* Beran 1994)

$$X_{t+l}(\mp) = \hat{X}_t(l) + z_p^{\mp}\left[1 + \sum_{j=1}^{l-1} \psi_j^2\right]^{1/2} \sigma_\varepsilon \tag{3.9}$$

The authors conjecture that the confidence intervals can be estimated more rigorously through calculating the standardized lower and upper bounds z_p^- and z_p^+ from the sample probability densities of the residuals. The residuals, as well as z_p^- and z_p^+, depend on the data signal and the model parameters in Eq. (2.2). Equation (3.10) below holds for the 95 % confidence limits, and one can write other confidence limits similarly.

$$P\left[z_p^- \le \frac{\hat{X}_{t+1} - X_t\,(l)}{\left[1 + \sum_{j=1}^{l-1} \psi_j^2\right]^{1/2} \sigma_\varepsilon} \le z_p^+\right] = 0.95 \tag{3.10}$$

The 2.5 and 97.5 percentiles of the residuals may be estimated to find z_p^- and z_p^+. If the residuals possess a normal distribution, then $z_p^{\mp} = \mp 1.96$.

One can infer from Eqs. (3.4) and (3.9) that the forecast confidence interval size depends on the variance and probability distribution of the residuals, forecast lead time *l*, the difference parameter d, and the autoregressive and the moving average coefficients Fig. (3.2).

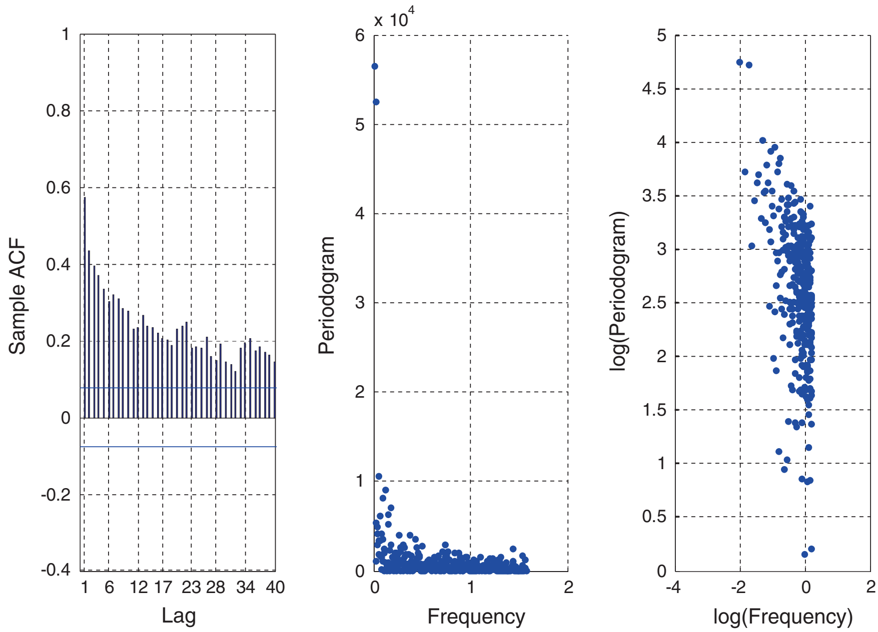

Fig. 3.2 Sample ACF, periodogram, and the logarithm of the periodogram of minimum water levels of Nile River

3.3 Updating

Updating of a forecast is important during a statistical forecasting process. However, it is usually neglected in forecast applications in the literature. Box and Jenkins (1976) provide the methodology for forecast updating for ARIMA models. Assuming that forecasts at origin t are available for lead times 1, 2, …L, then as soon as X_{t+1} becomes available, the forecast error $\varepsilon_{t+1} = X_{t+1} - \hat{X}_t(1)$ can be calculated, and one may update the forecasts using ψ weights (Box and Jenkins, 1976) as

$$\hat{X}_{t+1}(l) = \underset{t}{\hat{X}}(l+1) + \psi_l \varepsilon_{t+1} \tag{3.11}$$

for lead times 1, 2, …, L−1. Eq. (3.11) was originally developed for the ARIMA models, but in this study it is used for the ARFIMA models as well when the ψ weights are estimated from Eq. (3.6) for the non-integer differencing parameter d.

References

Box GEP, Jenkins GM (1976) Time series analysis: forecasting and control. Holden-Day, San Fransisco

Beran J (1994) Statistics for long-memory processes. Chapman and Hall, New York