Chapter 18
Time Series Analysis

18.1 Introduction

Time series analysis is the technique used to study observations that are measured over time. Examples include natural phenomena (temperature, humidity, wind speed) and business variables (price of commodities, stock market indices) that are measured at regular intervals (hourly, daily).

Like regression analysis, time series analysis seeks to model a response variable as a function of one or more explanatory variables. Time series analysis differs from other forms of regression analysis in one fundamental way. Previously we have assumed the observations are uncorrelated with each other, except perhaps through their dependency on the explanatory variables. Thus, in regression, we would model

$$Y = X\beta + \epsilon$$

and make the assumption that $\text{corr}(\epsilon_i, \epsilon_j) = 0$. In time series analysis we do not make the assumption of independence. Instead we make an explicit assumption of `dependence` and the task of the analysis is to model the dependence.

Conventionally, the notation used to denote a time series is $X_t, t = 1, 2, \ldots, n$. In this chapter we assume that the successive times at which observations are taken are equally spaced apart, for example, monthly, quarterly, or annual observations. (Extensions to non-equally spaced observations are in the **zoo** package in R.) Initially, we also assume that X_t is a *stationary* zero-mean time series. (A time series is said to be stationary if the distribution of $\{X_t, \ldots, X_{t+n}\}$ is the same as that of $\{X_{t+k}, \ldots, X_{t+n+k}\}$ for any choice of t, n, and k.) This means that as time passes, the series does not drift away from its mean value. Often when stationarity is absent it can be achieved by analyzing differences between successive terms of the time series rather than the original time series itself.

R.M. Heiberger, B. Holland, *Statistical Analysis and Data Display*,
Springer Texts in Statistics, DOI 10.1007/978-1-4939-2122-5_18

The term *lag* is used to describe earlier observations in a sequence. We indicate lagged observations with the *backshift* operator B, defined to mean

$$BX_t = B^1 X_t = X_{t-1}$$

and by extension

$$B^2 X_t = B(BX_t) = BX_{t-1} = X_{t-2}$$

Often B is used as the argument of a polynomial function. For example, if

$$\psi(B) = 3B^2 - 4B + 2$$

then

$$\psi(B)X_t = 3X_{t-2} - 4X_{t-1} + 2X_t$$

There is a distinction between the residual error ϵ in regression analysis and ε components of time series models. A time series $X_t, t = 1, 2, \ldots, n$, is a special case of a large class of models known as *stochastic processes*. Random variables $\varepsilon_t, \varepsilon_{t-1}, \ldots$ are *random shocks* to the process. This shock concept is distinct from regression residuals ϵ that represent the inability of model predictors to completely explain the response. In some time series models a linear combination of lagged ε's, $\sum \theta_k \varepsilon_{t-k}$, can be viewed as an error concept analogous to ϵ in regression.

The ARIMA class of models discussed in this chapter can be fit to most regular time series that exhibit systematic behavior with random perturbations that are small compared to the systematic components. They are not appropriate for modeling time series having irregular cyclical behavior (such as the business cycle when modeling Gross National Product) or irregular sizeable shocks (such as federal spending for relief from natural disasters such as major hurricanes, floods, earthquakes, etc.).

Standard time-related data manipulations are easier when the time parameter is built into the data object. Fundamental operations like comparisons of two series or merging two series (`ts.union` or `ts.intersect`) are easily specified and the program automatically aligns the time parameter. Table 18.1 illustrates the two alignment options.

18.2 The ARIMA Approach to Time Series Modeling

In this chapter we introduce the Box–Jenkins ARIMA approach to time series modeling Box and Jenkins (1976). This methodology involves two primary types of dependence structures, autoregression and moving averages, as well as the concept of differencing. We assume throughout that the independent random shocks ε_t are distributed with mean 0 and a common variance σ^2.

Table 18.1 Alignment of time series with the (`ts.union` or `ts.intersect`) functions.

```
> x <- ts(sample(10), start=1978)

> y <- ts(sample(6), start=1980)

> x
Time Series:
Start = 1978
End = 1987
Frequency = 1
 [1] 10  6  9  1  4  3  2  7  8  5

> y
Time Series:
Start = 1980
End = 1985
Frequency = 1
[1] 3 4 6 1 2 5

> ts.union(x,y)
Time Series:
Start = 1978
End = 1987
Frequency = 1
      x  y
1978 10 NA
1979  6 NA
1980  9  3
1981  1  4
1982  4  6
1983  3  1
1984  2  2
1985  7  5
1986  8 NA
1987  5 NA
```

```
> ts.intersect(x,y)
Time Series:
Start = 1980
End = 1985
Frequency = 1
     x y
1980 9 3
1981 1 4
1982 4 6
1983 3 1
1984 2 2
1985 7 5
```

18.2.1 AutoRegression (AR)

The equation describing the first-order autoregression model AR(1) is

$$X_t = \phi X_{t-1} + \varepsilon_t \tag{18.1}$$

Each observation X_t is correlated with the preceding observation (at lag=1) X_{t-1} and, to a lesser extent, with all earlier observations. In the AR(1) model, each observation X_t has correlation ϕ with the preceding (at lag=1) observation X_{t-1}. The correlation of X_t with X_{t-k} is

$$\text{corr}(X_t, X_{t-k}) = \phi^k, \quad k = 1, 2, \ldots \tag{18.2}$$

That is, the correlation decreases exponentially with the length of lag. For example,

$$\begin{aligned}\text{corr}(X_t, X_{t-1}) &= \text{corr}(\phi X_{t-1} + \varepsilon_t,\ X_{t-1}) \\ &= \phi\end{aligned}$$

and

$$\begin{aligned}\text{corr}(X_t, X_{t-2}) &= \text{corr}(\phi X_{t-1} + \varepsilon_t,\ X_{t-2}) \\ &= \text{corr}(\phi(\phi X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t,\ X_{t-2}) \\ &= \phi^2\,\text{corr}(X_{t-2},\ X_{t-2}) + \text{corr}(\phi\varepsilon_{t-1} + \varepsilon_t,\ X_{t-2}) \\ &= \phi^2 + 0\end{aligned}$$

The AR(1) model is further discussed in Section 18.5.2.

With p-order lags, the autoregression equation is written as

$$\Phi_p(B)X_t = \varepsilon_t \tag{18.3}$$

where

$$\Phi_p(B) = \phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$$

is a p^{th}-degree polynomial. This model is referred to as AR(p). The AR(1) model in Equation (18.1) is the special case where $\Phi_p(B) = \Phi_1(B) = 1 - \phi B$.

18.2.2 Moving Average (MA)

The equation describing the first-order moving average model MA(1) is

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1} \tag{18.4}$$

This model is called "moving average" because the right-hand side is a weighted moving average of the independent random shock ε_t at two adjacent time periods.

Each observation X_t in the MA(1) model is correlated with the preceding observation X_{t-1} and is uncorrelated with earlier observations. For example,

$$\text{corr}(X_t, X_{t-1}) = -\theta/(1 + \theta^2)$$

and

$$\begin{aligned}\text{corr}(X_t, X_{t-2}) &= \text{corr}(\varepsilon_t - \theta\varepsilon_{t-1},\ \varepsilon_{t-2} - \theta\varepsilon_{t-3}) \\ &= 0\end{aligned}$$

With q-order lags, the equation is written as

$$X_t = \Theta_q(B)\varepsilon_t \tag{18.5}$$

where

$$\Theta_q(B) = \theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$$

is a q^{th}-degree polynomial. This model is denoted MA(q). The MA(1) model in Equation (18.4) is the special case where $\Theta_q(B) = \Theta_1(B) = 1 - \theta B$. The MA(1) model is further discussed in Section 18.5.3.

18.2.3 Differencing

Differencing of order 1 is defined by

$$\nabla X_t = (1 - B)X_t = X_t - X_{t-1} \tag{18.6}$$

Simple models are written for the differenced data, for example,

$$\nabla X_t = \varepsilon_t - \theta \varepsilon_{t-1} \tag{18.7}$$

or, equivalently

$$X_t - X_{t-1} = \varepsilon_t - \theta \varepsilon_{t-1}$$

Model (18.7) is structurally the same as Model (18.4) in that it has the same right-hand side. The left-hand sides differ. Model (18.7) uses the differenced time series ∇X_t as its response variable where Model (18.4) used the observed variable X_t. Differencing removes nonstationarity in the mean. More complicated models involving higher-order differencing are denoted by a polynomial

$$\nabla^d(B) = (1 - B)^d$$

The interpretation is

$$\nabla^1(B)X_t = X_t - X_{t-1}$$

18.2.4 Autoregressive Integrated Moving Average (ARIMA)

We work with both AR(p) and MA(q) with lags greater than or equal to 1, and with a combined situation called ARIMA(p, d, q) (autoregressive integrated moving average). The term *integrated* means that we use the AR and MA techniques on *differenced* data. The general form of the ARIMA(p, d, q) model is

$$\Phi_p(B)\nabla^d X_t = \Theta_q(B)\varepsilon_t \tag{18.8}$$

where ε_t is a random shock with mean zero and $\text{var}(\varepsilon_t) = \sigma_\varepsilon^2$.

There are many important special cases.

ARIMA(1,0,0) = AR(1) model is in Equation (18.1).

ARIMA(0,0,1) = MA(1) model is in Equation (18.4).

ARIMA(0,1,0) is the first difference model in Equation (18.6).

ARIMA(0,1,1) model is shown in Equation (18.7).

ARIMA(1,1,1) model looks like

$$\begin{aligned} \Phi_1(B)\nabla X_t &= \Theta_1(B)\varepsilon_t \\ (1-\phi B)(1-B)X_t &= (1-\theta B)\varepsilon_t \\ (1-(1+\phi)B+\phi B^2)X_t &= (1-\theta B)\varepsilon_t \\ X_t-(1+\phi)X_{t-1}+\phi X_{t-2} &= \varepsilon_t-\theta\varepsilon_{t-1} \end{aligned}$$

ARIMA$(p,0,q)$ with $d=0$, hence no differencing, is also called an ARMA(p,q) model (autoregressive moving average)).

18.3 Autocorrelation

Two principal tools for studying time series are the autocorrelation function (ACF) and the partial autocorrelation function (PACF). The ACF assists in the diagnosis of MA models. The PACF is used in the diagnosis of AR models.

18.3.1 Autocorrelation Function (ACF)

The defining equation for the lag-k autocorrelation coefficient ρ_k is

$$\rho_k = \text{acf}(k) = \text{corr}(X_t, X_{t-k})$$

The discrete function $\{\rho_k\}$ indexed by the lag k is called the autocorrelation function of the series Z. The sample estimators $\{r_k\}$ are defined by

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$$

$$c_k = \frac{1}{n}\sum_{t=k+1}^{n} (X_t - \bar{X})(X_{t-k} - \bar{X}) \quad \text{autocovariance}$$

$$r_k = c_k/c_0 \quad \text{autocorrelation}$$

Note that the division is always by n.

The ACF for a time series $z = (x_1, \ldots, x_n)$ is calculated in R by

```
acf(z)
```

18.3.2 Partial Autocorrelation Function (PACF)

The defining equation for the PACF is

$$\phi_{kk} = \text{pacf}(k) = \text{corr}(X_t, X_{t-k} \mid X_{t-1}, X_{t-2}, \ldots, X_{t-(k-1)})$$

The sample estimators are defined by solving the Yule–Walker equations that hold for an AR(p) process (see Box and Jenkins (1976) for details). An illustrative (but not practical) estimator is shown in the R function in Table 18.2. The PACF is calculated in R by

```
acf(z, type="partial")
```

Note that $\rho_1 = \phi_{11}$, that is pacf(1) = acf(1), in mathematics notation. In R notation, the same statement is

```
acf(x)["1"]$acf == acf(x, type="partial")["1"]$acf
```

for a numeric vector `x`.

18.4 Analysis Steps

There are three main steps in time series analysis using the ARIMA models of the Box–Jenkins approach.

Identification: choice of the proper transformations to apply to the time series, consisting of variance-stabilizing transformations and of differencing. Determining the number of model parameters: d, the order of differencing; p, the number of autoregressive parameters; and q, the number of moving average parameters.

Table 18.2 Illustrative definition of the PACF. Do **NOT** use in actual calculations!

```
> ## This function illustrates the definition of the pacf.
> ## Do NOT use in actual calculations!
> ##
> ## my.pacf requires a detrended series, otherwise the answer is
> ## nonsense, as it starts losing precision after the first few lags.
>
> my.pacf <- function(z, k=2) {
+   z <- z - mean(z)
+   x <- ts.intersect(z, lag(z,-1))
+   if (k==1) return(cor(x[,1], x[,2]))
+   for (kk in 2:k) x <- ts.intersect(x, lag(z,-kk))
+   nr <- nrow(x)
+   nc <- ncol(x)
+   r1 <- lm(x[,1]  ~ -1 + x[,-c(1,nc)])$resid
+   r2 <- lm(x[,nc] ~ -1 + x[,-c(1,nc)])$resid
+   cor(r1,r2)
+ }

> my.pacf(ozone.subset, 2)
[1] -0.2665

> acf(ozone.subset, type="partial", plot=FALSE)$acf[2]
[1] -0.2583
```

Estimation: estimation of the parameters of the identified model, usually by maximum likelihood.

Diagnostics: verification that the estimated model and parameters do indeed capture the essence of the behavior of the data.

We offer these recommendations for interpreting sequence plots, ACF plots, and PACF plots. They are based on the Box–Jenkins methodology described in the texts by Box and Jenkins (1976) and Wei (1990).

1. Trends in the sequence plot must be removed by differencing. This is required before attempting to interpret the ACF and PACF plots. The interpretation below of ACF and PACF plots depends on stationarity.
2. No correlation—white noise
   The ACF and PACF are negligible at all lags.
3. AR(p)
   The ACF decays slowly.
   The PACF cuts off at lag p.

4. MA(q)
   The ACF cuts off at lag q.
   The PACF decays slowly.

5. ARMA(p, q)
   The ACF decays slowly from lag $\max(q - p, 0)$ on.
   The PACF decays slowly from lag $\max(p - q, 0)$ on.

   The orders p and q usually can't be read directly from these plots. Looking at the ACF and PACF plots for models with larger values of p and q can be helpful. The ESACF (extended sample autocorrelation function) (see, for example, Wei (1990), p. 128) can also be helpful.

Several additional tools are used to identify well-fitting models.

- The *Akaike information criterion* (AIC) for a particular model is defined as $-2(\ln L) + 2m$, where L is the model's loglikelihood and m is the number of parameters needed to estimate the model. Like the C_p statistic used to decide among multiple regression models, introduced in Equation (9.28), the AIC is the sum of a goodness-of-fit component and a penalty for lack of simplicity. Low values of AIC are preferred to large values.
- The *portmanteau goodness-of-fit test* for a particular model at lag ℓ is actually a collection of tests, one for each $k = 1, 2, \ldots, \ell$. Each of these individual tests is a test of the negligibility of the autocorrelations of the model residuals up to and including lag ℓ. In a well-fitting model, these hypotheses should be retained because such a model should have negligible autocorrelations. Therefore, for well-fitting models the p-values of these tests should not be small.
- The highest-order AR and MA parameters of well-fitting models are significantly different from zero, indicating that the corresponding model terms and terms of lower order are needed. Such significance is suggested by a corresponding t statistic that exceeds 2 in absolute value. For example, a model with $p = 2$ must have ϕ_2 significantly different from zero, but it is not essential that ϕ_1 be significantly nonzero.
- A well-fitting model has an estimated residual variance $\hat{\sigma}^2$ at least as small as those of competing models. The estimated residual variance and the AIC carry similar information. The AIC is usually preferred to the residual variance because the AIC includes a penalty for lack of simplicity and the residual variance does not.

In most situations these diagnostics will point to the same uniquely best model or subset of equivalently well-fitting models.

18.5 Some Algebraic Development, Including Forecasting

Usually, the ultimate purpose of finding a well-fitting time series model is the production of forecasts and forecast intervals h periods beyond the final observation of the existing series. While we have shown how to produce forecasts and intervals in R, here we provide a brief introduction to the algebra behind such forecasts. The algebra is intractable by hand for all but a few special cases.

18.5.1 The General ARIMA Model

The time series model for a 0-mean time series X_t, with $E(X_t) = 0$ and $\text{var}(X_t) = \sigma^2$, is

$$\phi(B)\nabla^d X_t = \theta(B)\,\varepsilon_t \tag{18.9}$$

One way to rewrite (18.9) is

$$\varepsilon_t = \theta^{-1}(B)\phi(B)\nabla^d X_t$$

Once the coefficients of ϕ and θ have been estimated by maximum likelihood, the fitted model is expressed in terms of the calculated residuals as

$$\hat{\varepsilon}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)\nabla^d X_t$$

where $\hat{\theta}(\cdot)$ and $\hat{\phi}(\cdot)$ are the polynomials in B after the coefficient estimates have been substituted into $\theta(\cdot)$ and $\phi(\cdot)$.

The calculated residuals $\hat{\varepsilon}_t$ will be used in many subsequent calculations.

The model (18.9) can also be rewritten as

$$\begin{aligned} X_t &= \phi^{-1}(B)\nabla^{-d}\,\theta(B)\,\varepsilon_t \\ &\overset{\text{def}}{=} \psi(B)\,\varepsilon_t \\ &= (1 + \psi_1 B + \ldots)\,\varepsilon_t \\ &= \varepsilon_t + \psi_1\varepsilon_{t-1} + \ldots + \psi_k\varepsilon_{t-k} + \ldots \end{aligned}$$

where $\psi(B)$ may have an infinite number of terms. The number of terms is finite with purely MA models (where $\psi = \theta$) and infinite when there are AR or differencing factors. In order that the model be stationary and invertible (that is, explicitly solvable for X_t), it is required that the roots of both polynomials $\phi(B)$ and $\psi(B)$ lie outside the unit circle. In addition $\phi(B)$ and $\theta(B)$ must have no roots in common. If the polynomials have common roots, these roots can be factored out.

The nonzero-mean case is essentially the same. Let the nonzero-mean time series be $Y_t = X_t + \mu$. We can subtract the mean from the observed Y_t-values to construct a 0-mean times series $X_t = Y_t - \mu$ and then proceed.

When we use the model for forecasting h steps ahead, we use the equation

$$\hat{X}_{t+h} = E(X_{t+h}|X_t, X_{t-1}, \ldots) = (\psi_h + \psi_{h+1}B + \ldots)\,\varepsilon_t$$

with forecast error

$$e_{t+h} = X_{t+h} - \hat{X}_{t+h} = \varepsilon_{t+h} + \psi_1\varepsilon_{t+h-1} + \ldots + \psi_{h-1}\varepsilon_{t+1}$$

and with variance of the forecast error

$$\text{var}(e_{t+h}) = \sigma^2(1 + \psi_1^2 + \psi_2^2 + \ldots + \psi_{h-1}^2)$$

The forecast error for h-step ahead forecasts, and its variance, have exactly h terms. The ε_t are uncorrelated. The forecast errors are correlated.

Probability limits for the forecasts are calculated as

$$\hat{X}_{t+h} \pm z_{\alpha/2}\,\hat{\sigma}\sqrt{1 + \psi_1^2 + \psi_2^2 + \ldots + \psi_{h-1}^2}$$

18.5.2 Special Case—The AR(1) Model

Starting from $X_2 = \phi X_1 + \varepsilon_2$, incrementing the subscripts on X_t, and then back-substituting [for example, $X_3 = \phi(\phi X_1 + \varepsilon_2) + \varepsilon_3$], we eventually get

$$X_{t+h} = \phi^h X_t + \phi^{h-1}\varepsilon_{t+1} + \ldots + \phi\varepsilon_{t+h-1} + \varepsilon_{t+h}$$

As a consequence, we take $\hat{X}_{t+h} = \hat{\phi}^h X_t$. Further,

$$\text{var}(X_{t+h}) = \sigma^2(1 + \phi^2 + \phi^4 + \ldots + \phi^{2(h-1)})$$

$$= \sigma^2\left(\frac{1 - \phi^{2h}}{1 - \phi^2}\right)$$

A $100(1 - \alpha)\%$ prediction interval for X_{t+h} is

$$\hat{X}_{t+h} \pm z_{\alpha/2}\hat{\sigma}\sqrt{\frac{1 - \hat{\phi}^{2h}}{1 - \hat{\phi}^2}}$$

18.5.3 Special Case—The MA(1) Model

Here we have $X_{t+1} = \varepsilon_{t+1} - \theta_1 \varepsilon_t$, and the general formulas simplify to

$$\begin{array}{lll} \hat{X}_{t+1} & = -\theta_1 \hat{\varepsilon}_t & \text{for } h = 1 \\ \hat{X}_{t+h} & = 0 & \text{for } h > 1 \\ \text{var}(\hat{X}_{t+1}) & = \sigma^2 & \text{for } h = 1 \\ \text{var}(\hat{X}_{t+h}) & = \sigma^2(1 + \theta_1^2) & \text{for } h > 1 \end{array}$$

In the MA(q) models the $\hat{\varepsilon}_{t+j}$-values are known for past observations (those for which $j \leq 0$), hence they appear in the prediction equations. A $100(1-\alpha)\%$ prediction interval for X_{t+h} is

$$\begin{array}{ll} \hat{X}_{t+1} \pm z_{\alpha/2}\hat{\sigma} & \text{for } h = 1 \\ \hat{X}_{t+h} \pm z_{\alpha/2}\hat{\sigma}\sqrt{1 + \theta_1^2} & \text{for } h > 1 \end{array}$$

18.6 Graphical Displays for Time Series Analysis

We present a number of graphical displays to facilitate the identification and model checking steps of ARIMA(p, d, q) modeling. Much of this material previously appeared in Heiberger and Teles (2002). We discuss an extension of these displays to model time series with seasonal components in Section 18.8. A general discussion of the features of these graphs appears in Section 18.A of this chapter's appendix.

Table 18.3 summarizes the nine achievable models formed by possible combinations of the number of AR parameters ($p = 0, 1, 2$) and MA parameters ($q = 0, 1, 2$). The appearance of the left-hand and right-hand side of the model equations is shown for each value of $(p, 0, q)$.

Figures 18.1 and 18.3 are examples of coordinated plots useful for identifying an ARIMA time series model.

Figure 18.1 contains a plot of the original time series along with its autocorrelation function and partial autocorrelation function. (Figure 18.2 is comparable to Figure 18.1 but for a differenced time series.)

The set of plots in Figure 18.3 consists of the residual ACF and PACF, the portmanteau goodness-of-fit test statistic (GOF), the standardized residuals, and the Akaike information criterion (AIC). The panels in the first four sets of plots are indexed by the number of ARMA parameters p and q. The AIC plot uses p and q as plotting variables. The orders of differencing and the orders of the autoregressive and moving average operators have been limited to $0 \leq p, d, q, \leq 2$. While this limitation is usually reasonable in practice, it is not inherent in the software.

Table 18.3 3×3 layout for the ARIMA$(p, 0, q)$ models. All the time series diagnostic plots and summary tabular data are constructed on this pattern. The rows give the number of AR parameters ($p = 0, 1, 2$) and the corresponding left-hand side of the model equation. The columns give the number of MA parameters ($q = 0, 1, 2$) and the corresponding right-hand side of the model equation. For example, the $(1, 1)$ cell of the array shows the information for the ARIMA$(1, 0, 1)$ model:

$$X_t - \phi_1 X_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

In all the displays we show, the differencing parameter d ($d = 0$ in this example) and the seasonal parameters $(P, D, Q)_s$ (if any) are held constant.

		Moving average model — Right-hand side		
Autoregression model		$q = 0$	$q = 1$	$q = 2$
p	Left-hand side	ε_t	$\varepsilon_t - \theta_1 \varepsilon_{t-1}$	$\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$
$p = 0$	X_t	$(0, 0, 0)$	$(0, 0, 1)$	$(0, 0, 2)$
$p = 1$	$X_t - \phi_1 X_{t-1}$	$(1, 0, 0)$	$(1, 0, 1)$	$(1, 0, 2)$
$p = 2$	$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2}$	$(2, 0, 0)$	$(2, 0, 1)$	$(2, 0, 2)$

Each set of nine panels is systematically structured in a 3×3 array indexed by the number of AR parameters p and MA parameters q. All nine panels in a set are scaled identically. Thus the reader can scan a row or column of the array of panels and see the effect of adding one more parameter to either the AR or MA side of the model.

The graphics are used to analyze the monthly El Nino data in file `data(elnino)` from NIST (2005). The *El Nino* effect is thought to be a driver of world-wide weather. The *southern oscillation* is a predictor of El Nino. It is defined as the sea level barometric pressure at Tahiti minus the sea level barometric pressure at the Darwin Islands. Repeated southern oscillation values below -1 essentially defines an El Nino. Figures 18.1 and 18.2 show the reported data y_t=`elnino` and the first differences $\nabla y_t \stackrel{\text{def}}{=} y_t - y_{t-1}$. The horizontal dashed lines on the ACF and PACF plots are the critical values for $\alpha = .05$ tests of the hypothesis, at each individual lag k, that the correlation coefficient is zero. Spikes on these plots that fall outside these horizontal boundaries suggest the possibility of a nonzero correlation.

Figure 18.1 suggests that successive months' southern oscillations are positively associated: $\text{corr}(y_t, y_{t-1}) \approx 0.65$. To address the positive association between successive months we analyze first differences in Figure 18.2; this Figure does not suggest a need for additional differencing and its ACF and PACF for the first differences shows systematic behavior: the ACF cuts off at lag 1 and the PACF decays slowly. This suggests an ARIMA(0,1,1) model for the original series.

Since the ACF and PACF show systematic behavior, we proceed to Figure 18.3, a collection of five sets of coordinated plots on a single page designed to facilitate identifying the best ARIMA$(p, 1, q)$ model based on fits of the nine models with

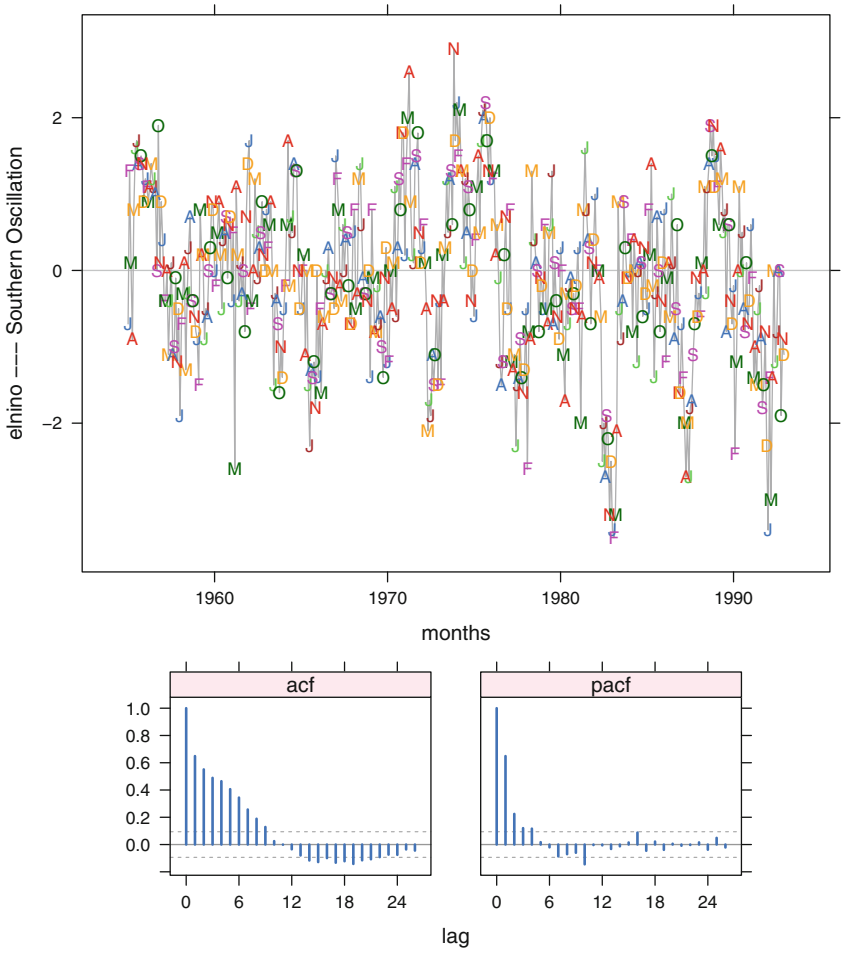

Fig. 18.1 Coordinated time series plot and ACF/PACF plots for the `elnino` time series: y_t. The response variable on the time series plot is the monthly value of the southern oscillation.

$0 \leq p, q \leq 2$. For each of these nine models, Figure 18.3 shows the ACF and PACF plots of the standardized residuals, a plot of the p-values of portmanteau goodness-of-fit tests at various lags, the Akaike information criterion arranged in the form of a pair of interaction plots, and the standardized residuals themselves. While in theory it is possible for p or q to exceed 2 (and the software permits larger values of p and/or q), this is unlikely to occur in practice provided that the data have been properly differenced and that any seasonal effects have been addressed.

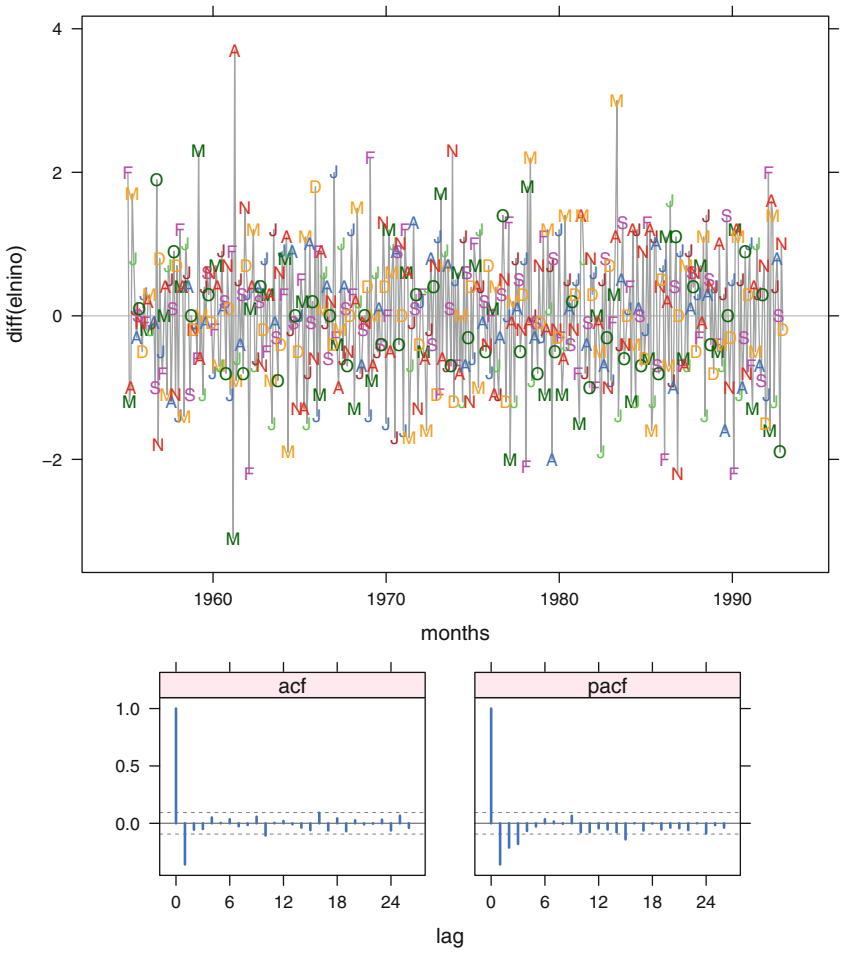

Fig. 18.2 Coordinated time series and ACF/PACF plots for the differenced `elnino` time series: ∇y_t.

Figure 18.3 confirms the choice $p = 0,\ q = 1$. For this parsimonious model,

- The ACF and PACF plots stay within the thresholds of significance.
- All p-values for the goodness-of-fit test exceed 0.05.
- The Akaike criterion is only slightly above that for all less parsimonious models.

Table 18.4 provides additional support for the ARIMA(0,1,1) model. The maximum likelihood estimate of the ARIMA parameters `coef` shows the MA(1) parameter to be 0.53. As this number is not close to 1, no further differencing is needed. The standardized value (denoted `t.coef`) of the MA(1) parameter is 13.2. This number greatly exceeds the usual critical value |2|.

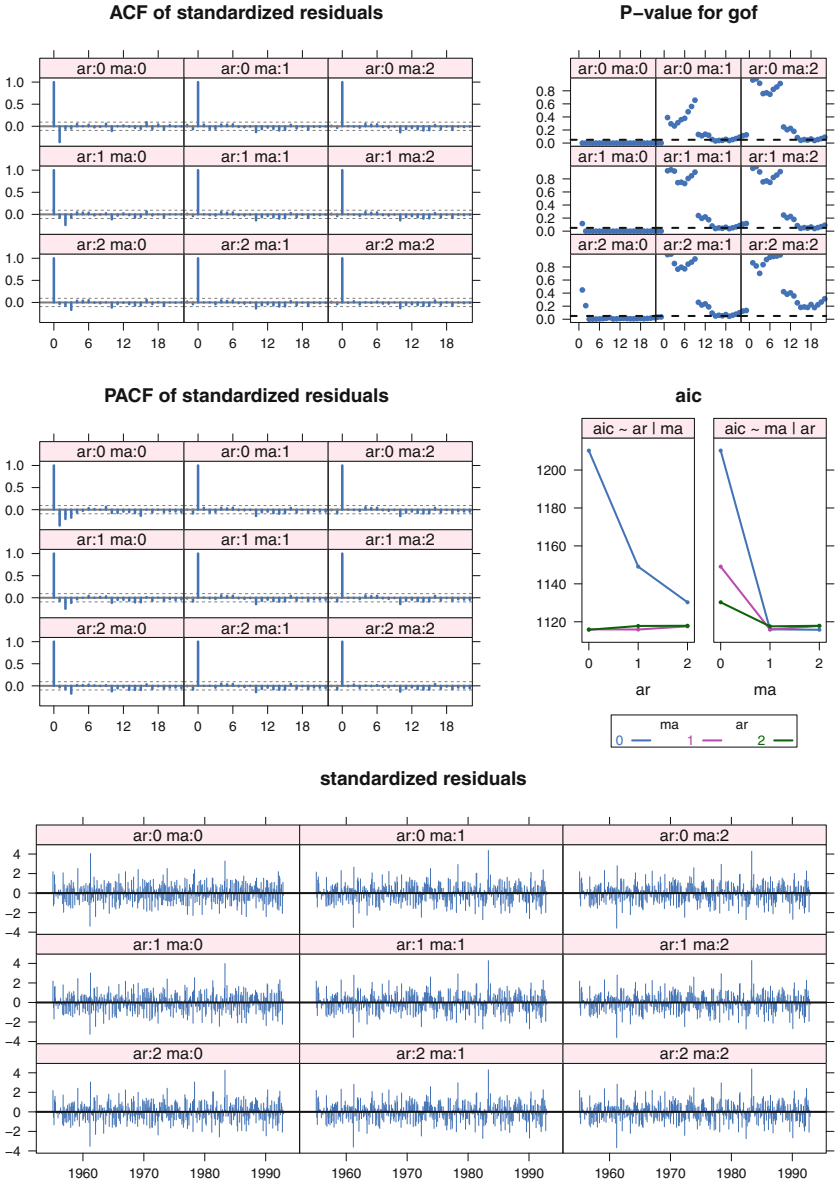

Fig. 18.3 Diagnostic plots for the set of models ARIMA(p,1,q) fit to the `elnino` data by maximum likelihood. Each set of nine panels is systematically structured in a 3×3 array with rows indexed by the number of AR parameters p and columns by the number of MA parameters q. All nine panels in a set are scaled identically. The AIC has been plotted as a pair of interaction plots: AIC plotted against q using line types defined by p; and AIC plotted against p, using line types defined by q.

Table 18.4 Estimation results for ARIMA($p, 1, q$) models fit to the elnino data.

```
> elnino.loop <- arma.loop(elnino, order=c(2,1,2))

> elnino.loop
$series
[1] "elnino"

$model
[1] "(p,1,q)"

$sigma2
       0      1      2
0 0.8333 0.6740 0.6707
1 0.7250 0.6708 0.6707
2 0.6925 0.6704 0.6678

$aic
     0    1    2
0 1210 1116 1116
1 1149 1116 1118
2 1130 1118 1118

$coef
             ar1      ar2     ma1      ma2
(0,1,0)       NA       NA      NA       NA
(1,1,0) -0.36153       NA      NA       NA
(2,1,0) -0.43720 -0.21245      NA       NA
(0,1,1)       NA       NA -0.5258       NA
(1,1,1)  0.12397       NA -0.6127       NA
(2,1,1)  0.07603 -0.04306 -0.5625       NA
(0,1,2)       NA       NA -0.4868 -0.06870
(1,1,2) -0.02671       NA -0.4604 -0.08237
(2,1,2)  0.73112 -0.23604 -1.2121  0.48199

$t.coef
            ar1     ar2     ma1     ma2
(0,1,0)      NA      NA      NA      NA
(1,1,0) -8.2319      NA      NA      NA
(2,1,0) -9.5122 -4.6147      NA      NA
(0,1,1)      NA      NA -12.492      NA
(1,1,1)  1.4328      NA  -8.944      NA
(2,1,1)  0.6434 -0.5906  -5.099      NA
(0,1,2)      NA      NA -10.391 -1.4911
(1,1,2) -0.0642      NA  -1.115 -0.3805
(2,1,2)  3.0449 -2.0250  -5.266  2.6781
```

18.7 Models with Seasonal Components

One of the strengths of the Box–Jenkins method is its handling of seasonal parameters. Time series frequently show seasonal patterns in their correlation structure. Most economic series have annual, quarterly, monthly, or weekly patterns as well as daily patterns. For example, retail sales figures often shows a surge in activity in December; power consumption figures show seasonal patterns as heating is used in the winter months and air conditioning in the summer months, and a weekly pattern where weekday consumption differs systematically from weekend consumption.

18.7.1 Multiplicative Seasonal ARIMA Models

Seasonal parameters are handled similarly to the nonseasonal parameters, with the subscripts varying by increments of the season. With monthly data, an annual season $s = 12$ is denoted by using 12-month lags, that is, X_t and X_{t-12}. We use uppercase Greek letters Φ and Θ to denote autoregressive and moving average polynomials, respectively, in the seasonal backshift operator B^s. The polynomials in the backshift B^s are denoted $\Phi(B^s)$ and $\Theta(B^s)$, and the differences are $\nabla_s = 1 - B^s$.

The seasonal portion of a seasonal model is denoted ARIMA$(P, D, Q)_s$ (with, for example, the seasonal $s = 12$ used for annual seasons when the underlying data is monthly, and $s = 7$ for weekly seasons when the underlying data is daily), where

P is the number of lags in the seasonal AR portion of the model, equivalently the order of the polynomial

$$\Phi(B^s) = 1 - \Phi_1 B^s - \ldots - \Phi_P B^{sP}$$

Q is the number of lags in the seasonal MA portion of the model, equivalently the order of the polynomial

$$\Theta(B^s) = 1 - \Theta_1 B^s - \ldots - \Theta_Q B^{sQ}$$

D is the number of seasonal differences prior to the AR and MA modeling, equivalently the power of the differencing binomial

$$\nabla_s^D = (1 - B^s)^D$$

The general multiplicative seasonal model, denoted

$$\text{ARIMA}(p, d, q) \times (P, D, Q)_s$$

is given by

$$\Phi(B^s)\,\phi(B)\,\nabla_s^D \nabla^d X_t = \Theta(B^s)\,\theta(B)\,\varepsilon_t \tag{18.10}$$

For various technical reasons, the roots of the seasonal polynomials $\Phi(B)$ and $\Psi(B)$ must satisfy certain conditions that parallel the restrictions on $\phi(B)$ and $\psi(B)$ mentioned in Section 18.5.1. The roots of $\Phi(B)$ and $\Psi(B)$ must lie outside the unit circle to assure that the model is stationary and invertible (solvable for $X(t)$). In addition, $\Phi(B)$ and $\Psi(B)$ must have no common roots. If these polynomials have common roots, these roots can be factored out.

*18.7.2 Example—*co2 *ARIMA*$(0,1,1)\times(0,1,1)_{12}$ *Model*

The final model for the co2 example discussed in Section 18.8 is written as ARIMA$(0,1,1)\times(0,1,1)_{12}$ model for X_t:

$$\nabla_{12}\nabla X_t = (1-\theta_1 B)\left(1-\Theta_1 B^{12}\right)\varepsilon_t \tag{18.11}$$

which expands to

$$X_t - X_{t-1} - X_{t-12} + X_{t-13} = \varepsilon_t - \theta_1\varepsilon_{t-1} - \Theta_1\varepsilon_{t-12} + \theta_1\Theta_1\varepsilon_{t-13}$$

18.7.3 Determining the Seasonal AR and MA Parameters

The procedure for determining the order P and Q of the seasonal parameters is comparable to the recommendations given in Section 18.4 for determining the order p and q of the nonseasonal parameters. As before, we work with the ACF and PACF for an appropriately differenced model. The distinction is that in examining the behavior of the ACF and PACF for seasonality, we examine only the values at seasonal intervals. For example, for monthly data with annual season ($s = 12$), these plots are examined at $t = 12, 24, 36, \ldots = 12\times(1,2,3,\ldots)$, ignoring values at other times. We then visualize the cutoff or decay behavior where *lag* now refers to seasonal intervals. If the ACF decays slowly at $t = 12, 24, 36, \ldots = 12\times(1,2,3,\ldots)$ and the PACF cuts off at $t = 24 = 12\times 2$, then $P = 2$ and $Q = 0$. If the PACF decays slowly at $t = 12, 24, 36, \ldots = 12\times(1,2,3,\ldots)$ and the ACF cuts off at $t = 12\times 1$, then $P = 0$ and $Q = 1$.

18.8 Example of a Seasonal Model—The Monthly `co2` Data

We extend the graphical displays discussed in Section 18.6 to the identification and model checking steps of ARIMA$(p, d, q) \times (P, D, Q)_s$ modeling. These graphs also first appeared in Heiberger and Teles (2002). A general discussion of the features of these graphs is deferred to Section 18.A.

The graphics are illustrated with one of the time series datasets distributed in the R package **datasets**, the Mauna Loa Carbon Dioxide Concentration series collected by the Scripps Institute of Oceanography, in La Jolla, California. The source is the climatology database maintained by the Oak Ridge National Laboratory Peterson (1990). These data represent monthly CO_2 concentrations in parts per million (ppm) from January 1959 to December 1990. Missing values have been filled in by linear interpolation.

Figures 18.4 through 18.7 are structured presentations of the plots of the series itself, of the ACF, and of the PACF. We show a magnified section of the plot for a five-year interval in Figure 18.5 Figure 18.4 displays the raw data series while Figure 18.6 displays the differenced series (monthly) and Figure 18.7 displays the twice-differenced series (monthly and annually).

18.8.1 Identification of the Model

Figure 18.4 is the plot of the observed data. The plot of the series itself shows a strong upward trend and a systematic labeling, with peaks occurring in the spring months and troughs in the autumn months. It is clear that the mean of this series is not constant over time. Both the ACF and PACF show systematic behavior. The ACF exhibits large values and a very slow decay with an annual periodicity. The PACF has large values and an annual periodicity. The conclusion is that the series is nonstationary, that is, it does not have a constant mean, and its autocorrelation function is time-dependent, implying that it shows nonrandom time-dependent behavior. Monthly differencing is required to model the nonstationarity, and annual differencing is necessary to remove the periodicity.

The time series and ACF and PACF plots for the differenced series ∇X_t in Figure 18.6 also show systematic annual behavior. The time series plot shows August/September troughs. The ACF exhibits a very slow decay at the seasonal lags, lags that are multiples of the seasonal period $s = 12$ months. This confirms that seasonal differencing with period 12 is required.

Figure 18.7 shows the time series (and the ACF and PACF) after non-seasonal and seasonal differencing $\nabla_{12}\nabla X_t$. There are no longer systematic components visible in the plot of the differences. The differenced series is stationary and it becomes

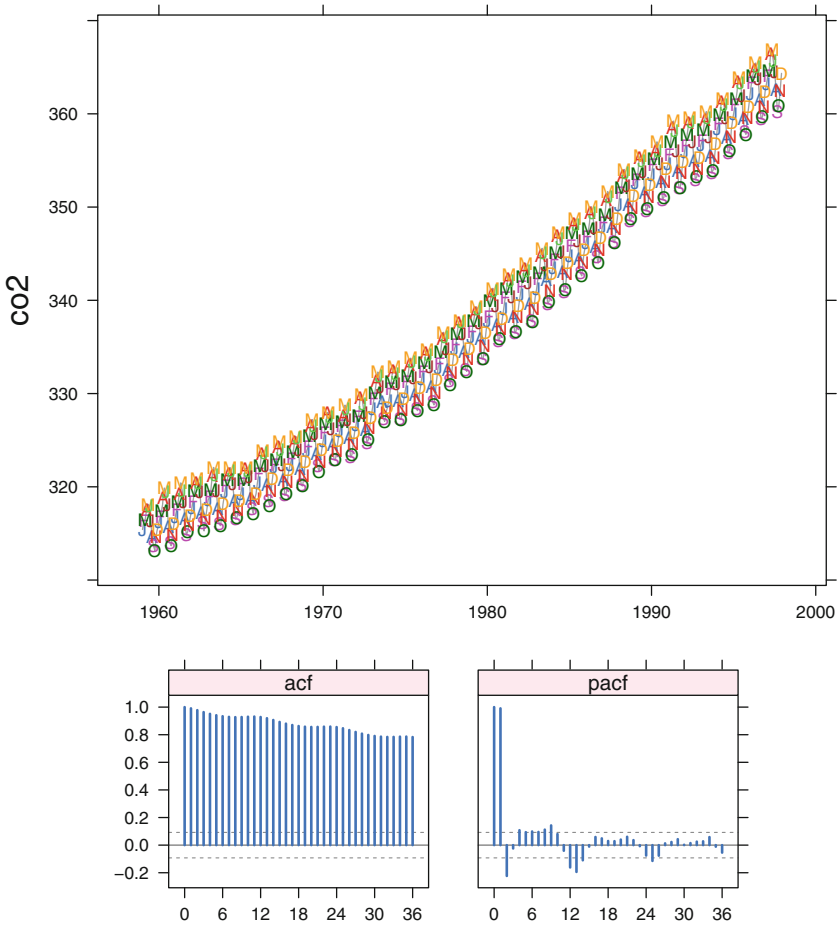

Fig. 18.4 Coordinated time series plot and ACF/PACF plots for the Mauna Loa CO_2 time series: X_t. The response variable on the time series plot is concentration in parts per million. We show a magnified section of the time series plot in Figure 18.5.

possible to identify a model for the series, that is, to look for the AR and MA parameters that best fit the twice-differenced data.

The nonseasonal component of the model of X_t is identified by looking at the first few monthly lags of the sample ACF and PACF of $\nabla_{12}\nabla X_t$ in Figure 18.7. The ACF seems to cut off after lag 1 and the PACF shows an exponential decay. The same type of behavior is seen at the seasonal lags, i.e., the ACF cuts off after lag 12 and the PACF shows an exponential decay at lags $12, 24, 36, \ldots$. These characteristics of the ACF and PACF suggest the ARIMA$(0, 1, 1) \times (0, 1, 1)_{12}$ model for X_t:

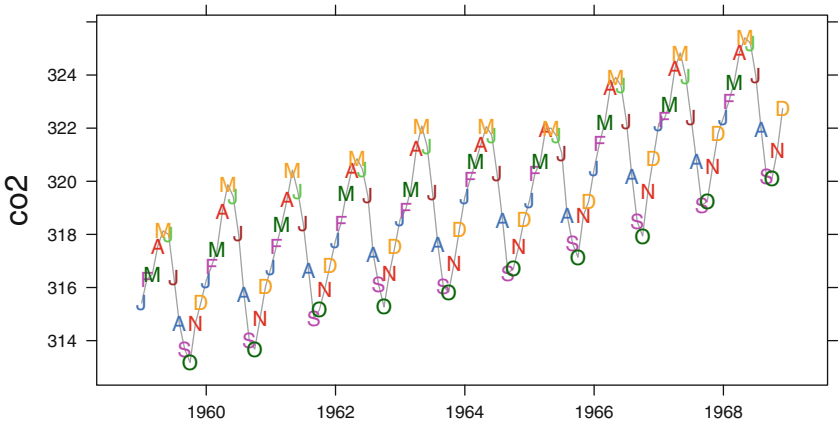

Fig. 18.5 Time series plot for 10 years of the Mauna Loa CO_2 time series: X_t. The data from 1959 to 1997 is shown in Figure 18.4.

$$\nabla_{12}\nabla X_t = (1 - \theta_1 B)\left(1 - \Theta_1 B^{12}\right)\varepsilon_t \tag{18.12}$$

A closer look at the ACF in Figure 18.7 indicates that it too may show an exponential decay in the first lags, suggesting that the ARIMA$(1,1,1)\times(0,1,1)_{12}$ model

$$(1 - \phi_1 B)\nabla_{12}\nabla X_t = (1 - \theta_1 B)\left(1 - \Theta_1 B^{12}\right)\varepsilon_t \tag{18.13}$$

might also be appropriate.

18.8.2 Parameter Estimation and Diagnostic Checking

In general, when analyzing seasonal time series data, initial guesses of at least some of the parameters p, q, P, Q may be provided from inspections of coordinated plots of original and differenced data such as Figures 18.4–18.7. Figures 18.8 and 18.9 each simultaneously consider nine models produced with the user function `arma.loop` described in Section 18.A.4. Figures in this class can be used to suggest seasonal parameters P and Q for a given set of nonseasonal and differencing parameters p, q, d, D, or to suggest nonseasonal parameters p and q for a given set of seasonal and differencing parameters P, Q, d, D. Alternating consideration of figures of both of these types can be used to settle on a final model.

Continuing with the `co2` data, Figure 18.8 displays a set of diagnostic plots for several models without a seasonal component, the ARIMA$(p,1,q)\times(0,1,0)_{12}$ models with $0 \le p, q \le 2$, that have been fit to the series $\nabla_{12}\nabla X_t$.

Since the `co2` data exhibit a seasonal behavior, Figure 18.8 is expected to confirm that seasonal parameters are required in the model of $\nabla_{12}\nabla X_t$. All the residual ACF

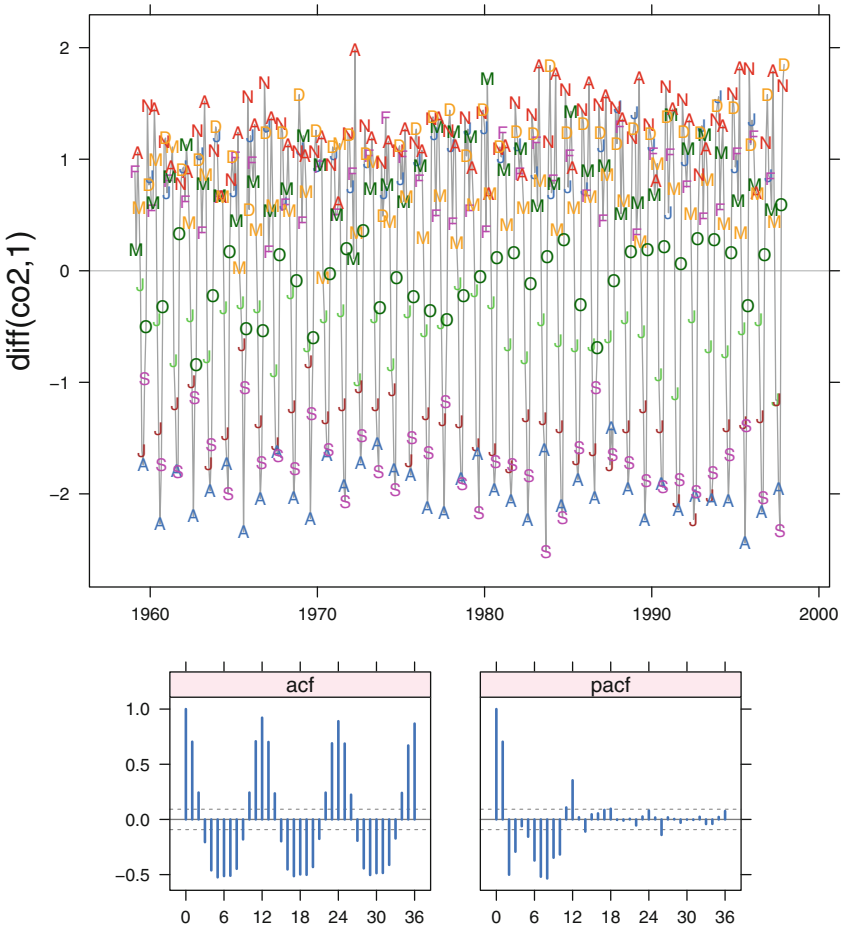

Fig. 18.6 Coordinated time series plot and ACF/PACF plots for the differenced Mauna Loa CO_2 time series: ∇X_t.

plots show a significant spike at `lag=12` months, and all the GOF plots show a break at the same `lag=12` months.

The residual ACF, PACF, and GOF plots in Figure 18.8 clearly confirm that seasonal parameters are necessary. The cutoff after the spike at `lag=12` of the residual ACF, and the exponential decay of the residual PACF at the seasonal lags (those that are multiples of 12 months), show that a seasonal MA parameter is necessary. This agrees with the identification of candidate models (18.11) and (18.13).

Next consider the ARIMA$(p, 1, q) \times (0, 1, 1)_{12}$ models with $0 \leq p, q \leq 2$. The diagnostic plots for models including the seasonal MA parameter are in Figure 18.9.

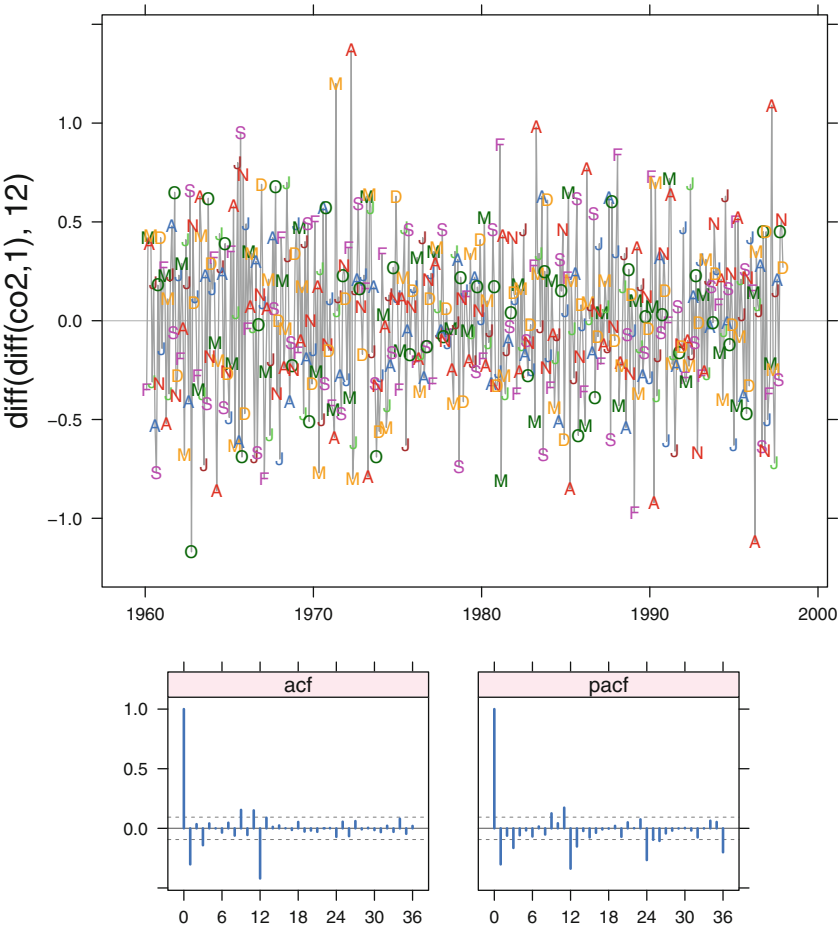

Fig. 18.7 Coordinated time series plot and ACF/PACF plots for the twice-differenced Mauna Loa CO_2 time series: $\nabla_{12}\nabla X_t$.

The $q = 0$ column of the residual ACF and GOF plots shows poor fits. The $q = 1$ column appears better than the $q = 2$ column. All three GOF plots for $q = 1$ are similar. The AIC plots show almost identical values when $q = 1$. This is seen as three almost coincident points at $q = 1$ in the "`aic ~ ma | ar`" plot and as a horizontal line for $q = 1$ over all three values of p in the "`aic ~ ar | ma`" plot. The conclusion is that one nonseasonal MA parameter is necessary.

Table 18.5 shows the AIC, the estimates of σ_ε^2, and the estimates of the ARMA parameters with their t-statistics for the set of ARIMA$(p, 1, q) \times (0, 1, 1)_{12}$ models. From the `t.coef` section of Table 18.5, the t statistics for both AR parameters in the

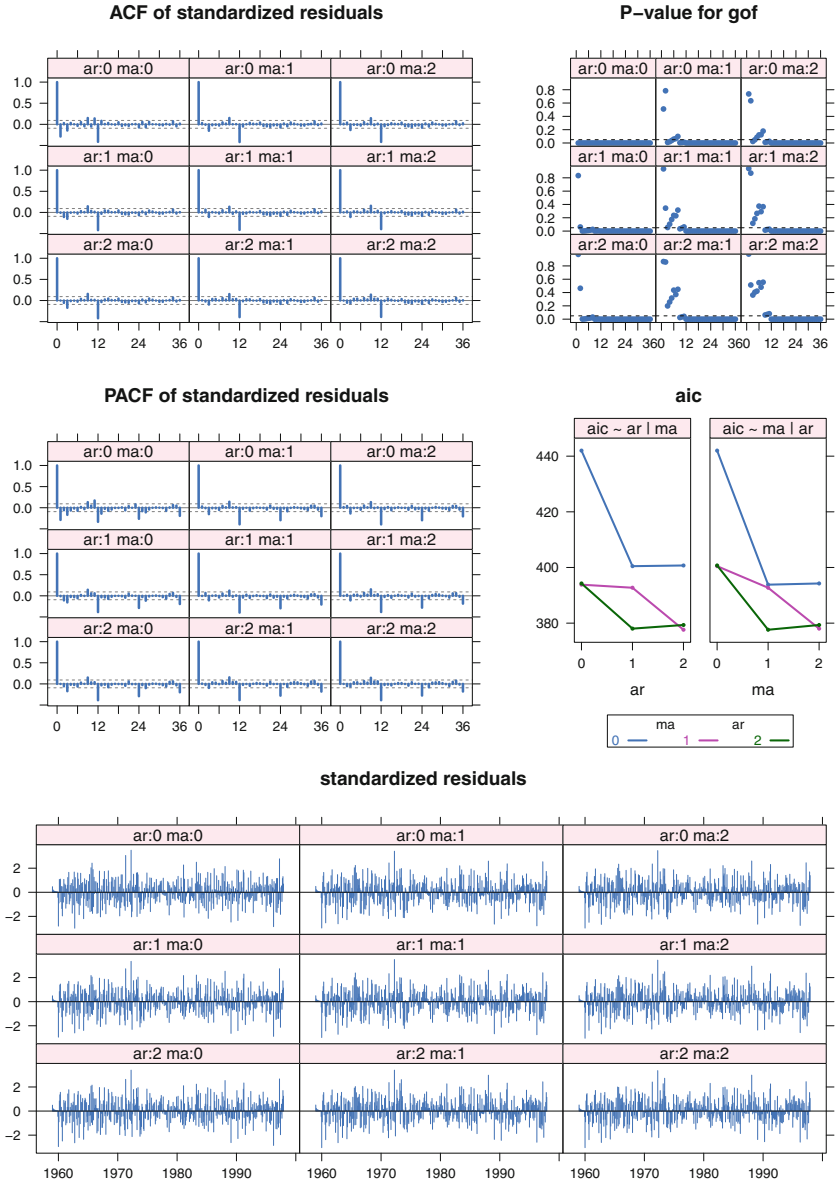

Fig. 18.8 Diagnostic plots for the set of models ARIMA$(p, 1, q) \times (0, 1, 0)_{12}$ fit to the CO_2 data by maximum likelihood. Each set of nine panels is systematically structured in a 3×3 array with rows indexed by the number of AR parameters p and columns by the number of MA parameters q. All nine panels in a set are scaled identically. The AIC is plotted as a pair of interaction plots: AIC plotted against q using line types defined by p; and AIC plotted against p, using line types defined by q.

ARIMA$(2,1,1)\times(0,1,1)_{12}$ model are not significant and this model can be rejected. The t-statistic for the AR(1) parameter in the ARIMA$(1,1,1)\times(0,1,1)_{12}$ model is marginally significant, leading us to consider both the models ARIMA$(0,1,1)\times(0,1,1)_{12}$ and ARIMA$(1,1,1)\times(0,1,1)_{12}$ for X_t. A different criterion is needed to distinguish between them. Both models are consistent with the analysis at the identification stage.

The detailed display of the estimation results for the ARIMA$(0,1,1)\times(0,1,1)_{12}$ model is shown in Table 18.6 (as displayed by the new print method for `arima` objects). A similar display for the ARIMA$(1,1,1)\times(0,1,1)_{12}$ model in Table 18.7 shows that the AR(1) and MA(1) parameters are highly correlated ($r = -.0167184/\sqrt{.02040605\times.0152966} = -0.9463$). The ARIMA$(1,1,1)\times(0,1,1)_{12}$ models can be discarded from further consideration.

The final step is the verification of the adequacy of the ARIMA$(0,1,1)\times(0,1,1)_{12}$ model. The residual ACF and PACF plots exhibit no significant spikes and all the GOF p-values are also not significant, showing that the residuals are approximately white noise. The AIC values have dropped from 310 in Figure 18.8 to 136 in Figure 18.9. The standardized residuals in Figure 18.9 are not inconsistent with the normal distribution. The ARIMA$(0,1,1)\times(0,1,1)_{12}$ model seems to be appropriate for X_t and the estimated model is

$$\begin{aligned}\nabla_{12}\nabla X_t &= \left(1-\hat{\theta}_1 B\right)\left(1-\hat{\Theta}_1 B^{12}\right)\hat{\varepsilon}_t \\ &= (1-0.36338\,B)(1-0.85806\,B^{12})\,\hat{\varepsilon}_t \\ \hat{\sigma}^2_{\varepsilon} &= 0.080299\end{aligned} \tag{18.14}$$

Table 18.5 Estimation results for ARIMA$(p, 1, q) \times (0, 1, 1)_{12}$ models fit to the CO_2 data.

```
> ddco2.loopPQ <-
+        arma.loop(co2,
+                  order=c(2,1,2),
+                  seasonal=list(order=c(0,1,1), period=12))

> ddco2.loopPQ
$series
[1] "co2"

$model
[1] "(p,1,q)x(0,1,1)12"

$sigma2
        0       1       2
0 0.09063 0.08260 0.08242
1 0.08358 0.08221 0.08214
2 0.08315 0.08176 0.08162

$aic
      0     1     2
0 221.5 178.2 179.1
1 184.2 178.1 180.1
2 183.5 177.8 179.1

$coef
                        ar1      ar2     ma1     ma2    sma1
(0,1,0)x(0,1,1)12        NA       NA      NA      NA -0.8887
(1,1,0)x(0,1,1)12 -0.292654       NA      NA      NA -0.8603
(2,1,0)x(0,1,1)12 -0.317119 -0.07825      NA      NA -0.8551
(0,1,1)x(0,1,1)12        NA       NA -0.3501      NA -0.8506
(1,1,1)x(0,1,1)12  0.239889       NA -0.5710      NA -0.8516
(2,1,1)x(0,1,1)12  0.390518  0.10540 -0.7329      NA -0.8544
(0,1,2)x(0,1,1)12        NA       NA -0.3436 -0.0492 -0.8499
(1,1,2)x(0,1,1)12 -0.962631       NA  0.6204 -0.3571 -0.8440
(2,1,2)x(0,1,1)12  0.007095  0.23191 -0.3477 -0.2473 -0.8548

$t.coef
                        ar1    ar2     ma1    ma2   sma1
(0,1,0)x(0,1,1)12        NA     NA      NA     NA -36.66
(1,1,0)x(0,1,1)12  -6.43561     NA      NA     NA -34.04
(2,1,0)x(0,1,1)12  -6.64989 -1.650      NA     NA -33.66
(0,1,1)x(0,1,1)12        NA     NA -7.0529     NA -33.15
(1,1,1)x(0,1,1)12   1.67707     NA -4.6167     NA -33.29
(2,1,1)x(0,1,1)12   3.01877  1.503 -6.2632     NA -33.52
(0,1,2)x(0,1,1)12        NA     NA -7.2649 -1.036 -33.26
(1,1,2)x(0,1,1)12 -37.32018     NA 11.4769 -7.224 -31.34
(2,1,2)x(0,1,1)12   0.01896  1.883 -0.9297 -1.150 -33.30
```

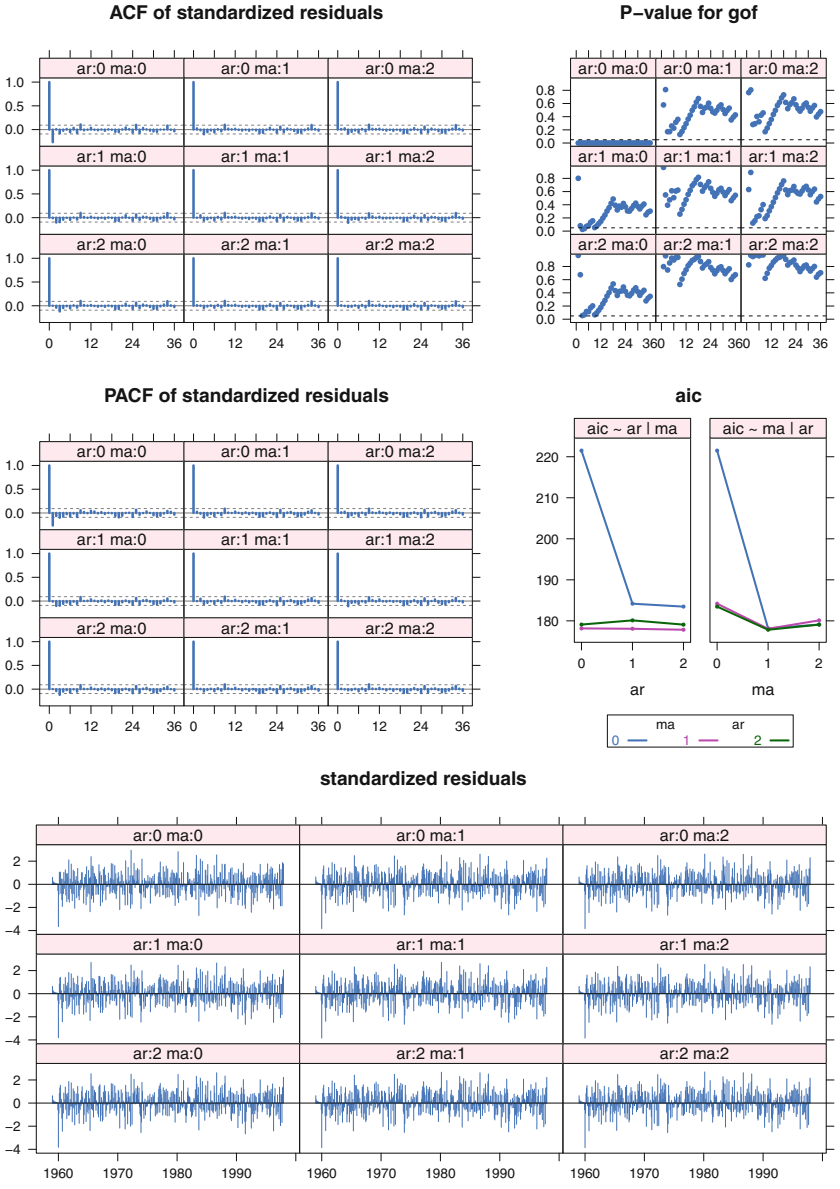

Fig. 18.9 Diagnostic plots for the set of models ARIMA$(p, 1, q) \times (0, 1, 1)_{12}$ fit to the CO_2 data by maximum likelihood.

Table 18.6 Estimation results for ARIMA$(0, 1, 1) \times (0, 1, 1)_{12}$ models fit to the CO_2 data.

```
> co2.arima <- ddco2.loopPQ[["0","1"]]

> co2.coef.t <- co2.arima$coef / sqrt(diag(co2.arima$var.coef))

> co2.arima

Call:
arima(x = x, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1),
      period = 12))

Coefficients:
        ma1    sma1
      -0.35  -0.851
s.e.   0.05   0.026

sigma^2 estimated as 0.0826:  log likelihood = -86.08,  aic = 178.2

> co2.coef.t
    ma1    sma1
 -7.053 -33.151

> vcov(co2.arima)
             ma1       sma1
ma1   0.0024638 -0.0002662
sma1 -0.0002662  0.0006583
```

Table 18.7 Estimation results for ARIMA$(1,1,1)\times(0,1,1)_{12}$ models fit to the CO_2 data.

```
> co2.arima11 <- ddco2.loopPQ[["1","1"]]

> co2.coef11.t <- co2.arima11$coef / sqrt(diag(co2.arima11$var.coef))

> co2.arima11

Call:
arima(x = x, order = c(1, 1, 1), seasonal = list(order = c(0, 1, 1),
      period = 12))

Coefficients:
        ar1     ma1    sma1
      0.240  -0.571  -0.852
s.e.  0.143   0.124   0.026

sigma^2 estimated as 0.0822:  log likelihood = -85.03,  aic = 178.1

> co2.coef11.t
    ar1     ma1    sma1
  1.677  -4.617 -33.287

> vcov(co2.arima11)
            ar1        ma1       sma1
ar1   0.0204605 -0.0167184 -0.0004893
ma1  -0.0167184  0.0152966  0.0002301
sma1 -0.0004893  0.0002301  0.0006544
```

18.8.3 Forecasting

The final plot in Figure 18.10 shows the last year of observed data and the forecasts, with their 95% forecast limits, obtained from the fitted model for the following year, i.e., for the months January through December 1998.

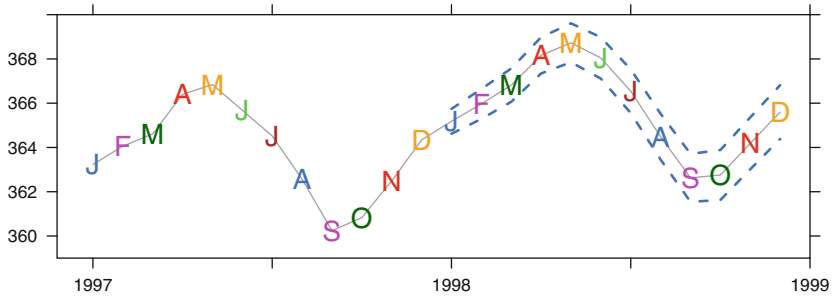

Fig. 18.10 CO_2—1997 observed, 1998 forecast + 95% CI.

18.9 Exercises

Many of the time series exercises ask you to construct and/or interpret plots of the time series itself, of the ACF and PACF, and of the diagnostics from a 3×3 set of ARIMA models. For Exercises 18.1, 18.2, and 18.3, go through this set of steps:

a. Describe the plot of the data and the ACF and PACF plots. Comment on whether you see anything systematic in the plot of the data. Are there spikes in the ACF and PACF plots. At which lags do they appear and what do they suggest? Do the ACF and PACF plots show any indication of a seasonal effect?

b. We chose to investigate a family of ARIMA$(p, 0, q)$ models, with $0 \le p, q \le 2$. Study the figures showing the diagnostic plots and the tables listing the parameter estimates. Describe each of the four sections of the diagnostic plot. What characteristics of each suggest a final model? Does the σ^2 (`sigma2`) section in the tables also suggest the same final model? Note for these three exercises, all of which are stationary and have zero mean, that the $(0, 0)$ panels of the ACF, PACF, and standardized residuals plots are essentially the same as the ACF, PACF, and time series plot of the data.

c. We printed the detail for the ARIMA$(1, 0, 1)$ model. Compare the ARIMA$(1, 0, 1)$ model to a simpler model with the closest σ^2. How do the AIC and the σ^2 compare? Would you recommend the simpler model? Why or why not?

18.1. Figure 18.11 shows the sequence, ACF, and PACF plots for a mystery time series X `data(tser.mystery.X)`. Figure 18.12 and Table 18.8 show the diagnostics and estimated coefficients obtained by fitting the 3×3 set of ARIMA$(p, 0, q)$ models to the series. Table 18.9 shows the detail for the ARIMA$(1, 0, 1)$ model. Study the graphs and tables and explain why and how they indicate that one of these models seems better suited to explain the data than the others.

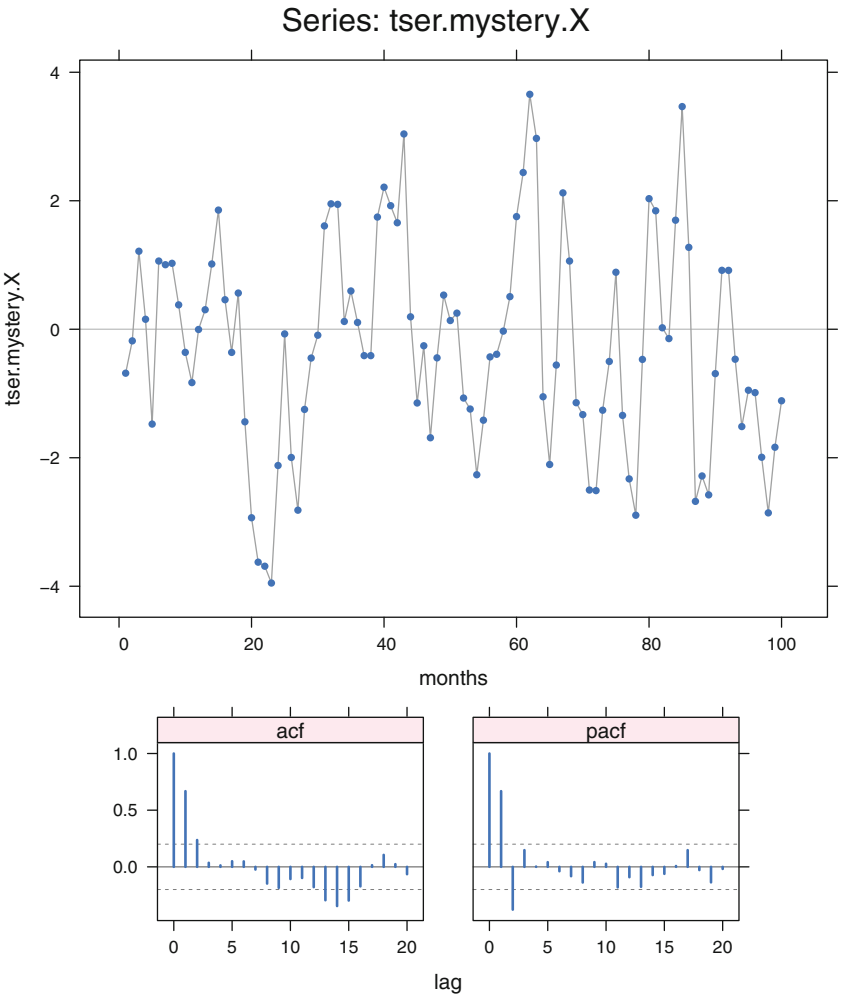

Fig. 18.11 Mystery time series X.

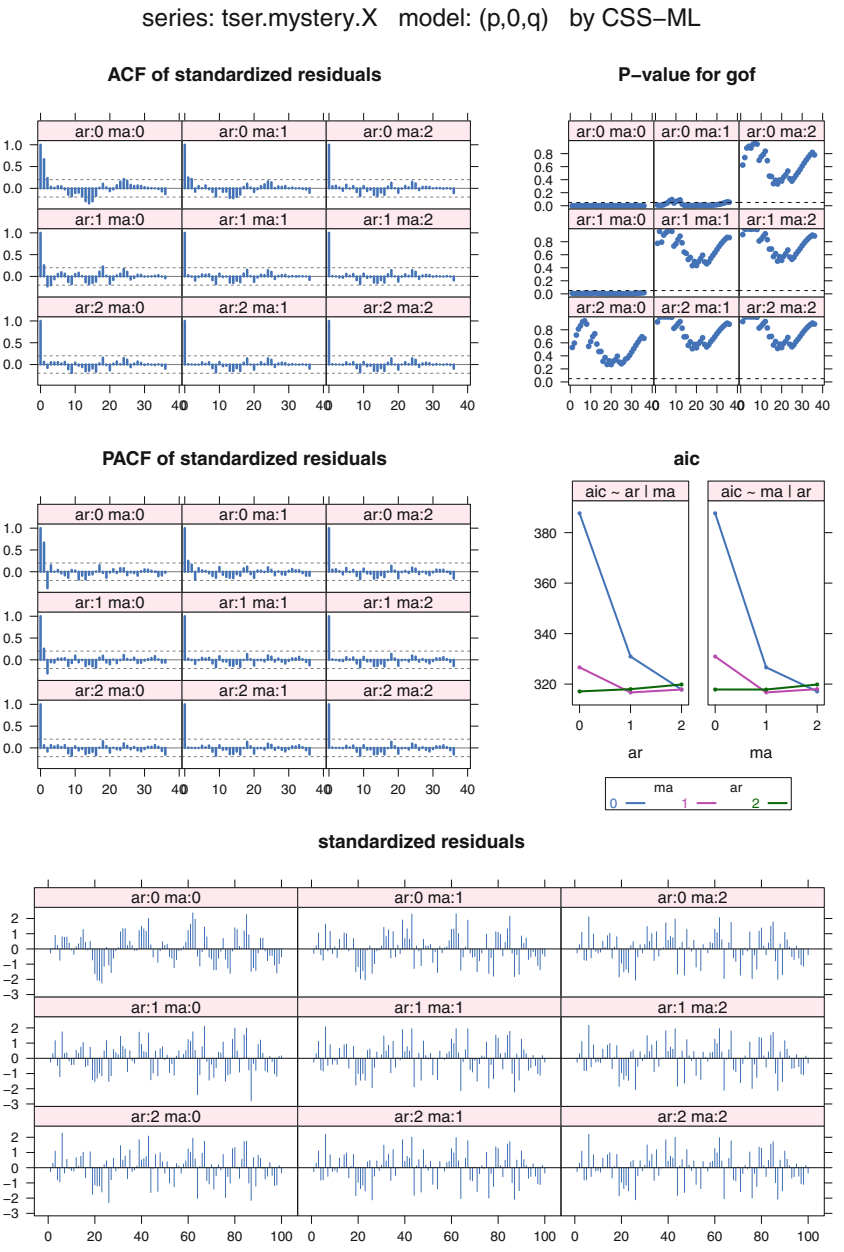

Fig. 18.12 Mystery time series X.

Table 18.8 Mystery time series *X*.

```
> X.loop
$series
[1] "tser.mystery.X"

$model
[1] "(p,0,q)"

$sigma2
      0     1     2
0 2.715 1.435 1.277
1 1.500 1.271 1.262
2 1.286 1.260 1.260

$aic
      0     1     2
0 387.7 326.6 317.1
1 330.9 316.7 318.0
2 317.9 317.9 319.8

$coef
           ar1     ar2    ma1     ma2 intercept
(0,0,0)     NA      NA     NA      NA        NA
(1,0,0) 0.6636      NA     NA      NA   -0.2744
(2,0,0) 0.9135 -0.3721     NA      NA   -0.2528
(0,0,1)     NA      NA 0.7264      NA   -0.2568
(1,0,1) 0.4299      NA 0.5221      NA   -0.2629
(2,0,1) 0.6213 -0.1781 0.3483      NA   -0.2571
(0,0,2)     NA      NA 0.9272 0.32958   -0.2548
(1,0,2) 0.2614      NA 0.7057 0.17341   -0.2588
(2,0,2) 0.5417 -0.1479 0.4280 0.04826   -0.2572

$t.coef
           ar1     ar2     ma1    ma2 intercept
(0,0,0)     NA      NA      NA     NA        NA
(1,0,0) 9.0295      NA      NA     NA   -0.7684
(2,0,0) 9.9262 -4.0464      NA     NA   -1.0257
(0,0,1)     NA      NA 13.2037     NA   -1.2471
(1,0,1) 3.8606      NA  5.1540     NA   -0.8828
(2,0,1) 2.6785 -0.9701  1.5262     NA   -0.9525
(0,0,2)     NA      NA 10.6280 3.5060   -1.0063
(1,0,2) 1.0949      NA  3.0019 0.8897   -0.9135
(2,0,2) 0.7891 -0.4658  0.6246 0.1267   -0.9478
```

Table 18.9 Mystery time series *X*.

```
> X.loop[["1","1"]]

Call:
arima(x = x, order = c(1, 0, 1))

Coefficients:
        ar1    ma1  intercept
      0.430  0.522     -0.263
s.e.  0.111  0.101      0.298

sigma^2 estimated as 1.27:  log likelihood = -154.3,  aic = 316.7
```

18.2. Figure 18.13 shows the sequence, ACF, and PACF plots for a mystery time series Y `data(tser.mystery.Y)`. Figure 18.14 and Table 18.10 show the diagnostics and estimated coefficients obtained by fitting the 3×3 set of ARIMA$(p, 0, q)$ models to the series. Table 18.11 shows the detail for the ARIMA$(1, 0, 1)$ model. Study the graphs and tables and explain why and how they indicate that one of these models seems better suited to explain the data than the others.

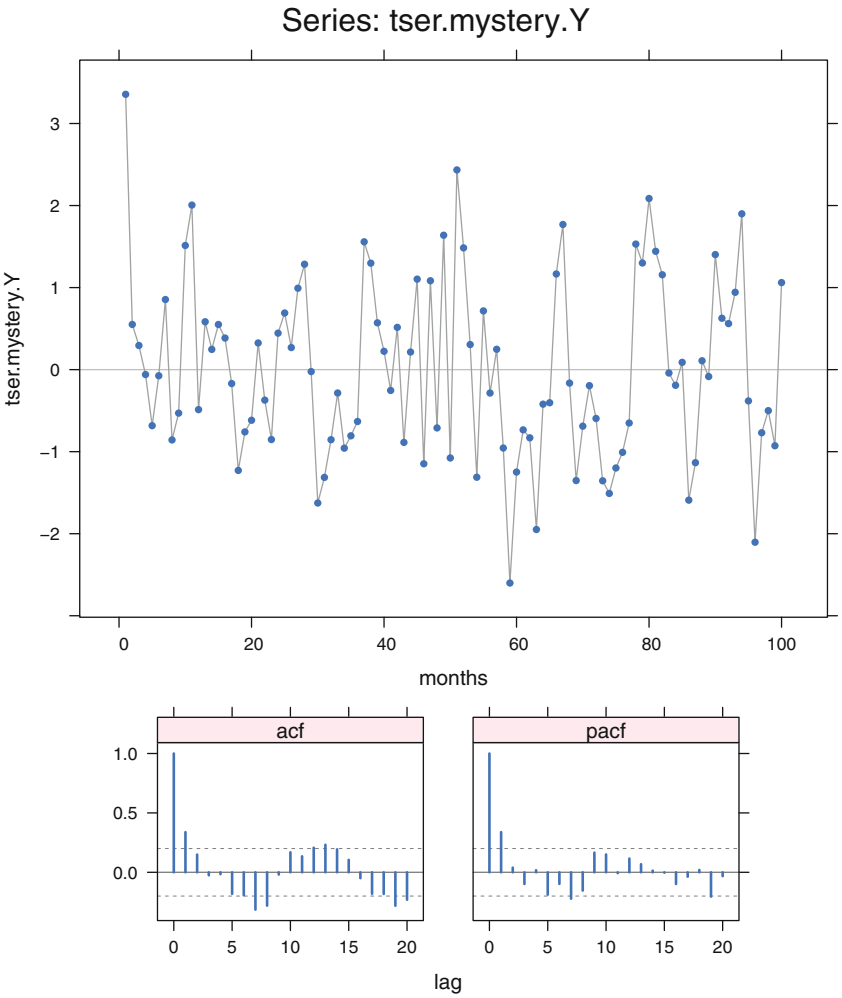

Fig. 18.13 Mystery time series Y.

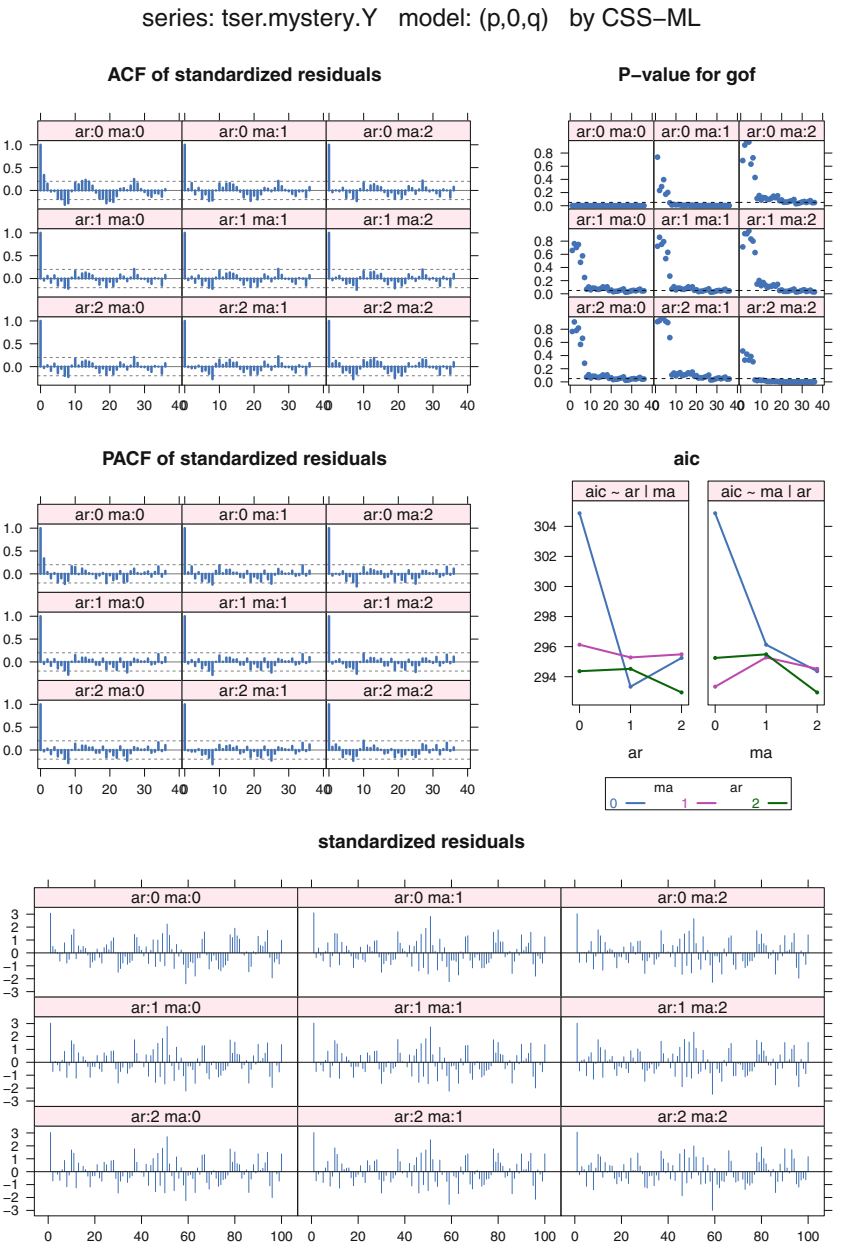

Fig. 18.14 Mystery time series Y.

Table 18.10 Mystery time series Y.

```
> Y.loop
$series
[1] "tser.mystery.Y"

$model
[1] "(p,0,q)"

$sigma2
      0     1      2
0 1.186 1.064 1.0243
1 1.035 1.034 1.0042
2 1.034 1.015 0.9502

$aic
      0     1     2
0 304.9 296.1 294.4
1 293.3 295.3 294.5
2 295.2 295.5 293.0

$coef
            ar1      ar2      ma1    ma2 intercept
(0,0,0)      NA       NA       NA     NA        NA
(1,0,0)  0.3740       NA       NA     NA   0.03050
(2,0,0)  0.3623  0.03129       NA     NA   0.03248
(0,0,1)      NA       NA  0.29684     NA   0.01777
(1,0,1)  0.4189       NA -0.05195     NA   0.03180
(2,0,1) -0.4676  0.38428  0.82758     NA   0.03363
(0,0,2)      NA       NA  0.37855 0.1854   0.02912
(1,0,2) -0.6314       NA  1.03122 0.4139   0.02815
(2,0,2) -0.9648 -0.54982  1.32535 0.9453   0.01509

$t.coef
           ar1     ar2     ma1    ma2 intercept
(0,0,0)     NA      NA      NA     NA        NA
(1,0,0)  3.817      NA      NA     NA    0.1884
(2,0,0)  3.433  0.2962      NA     NA    0.1944
(0,0,1)     NA      NA  3.6035     NA    0.1331
(1,0,1)  1.994      NA -0.2342     NA    0.1924
(2,0,1) -1.913  3.6606  3.3770     NA    0.1988
(0,0,2)     NA      NA  3.5048  2.022    0.1846
(1,0,2) -2.958      NA  5.0593  3.723    0.1879
(2,0,2) -8.179 -5.0155 21.2693 14.542    0.1191
```

Table 18.11 Mystery time series Y.

```
> Y.loop[["1","1"]]

Call:
arima(x = x, order = c(1, 0, 1))

Coefficients:
         ar1      ma1  intercept
      0.419  -0.052      0.032
s.e.  0.210   0.222      0.165

sigma^2 estimated as 1.03:  log likelihood = -143.6,  aic = 295.3
```

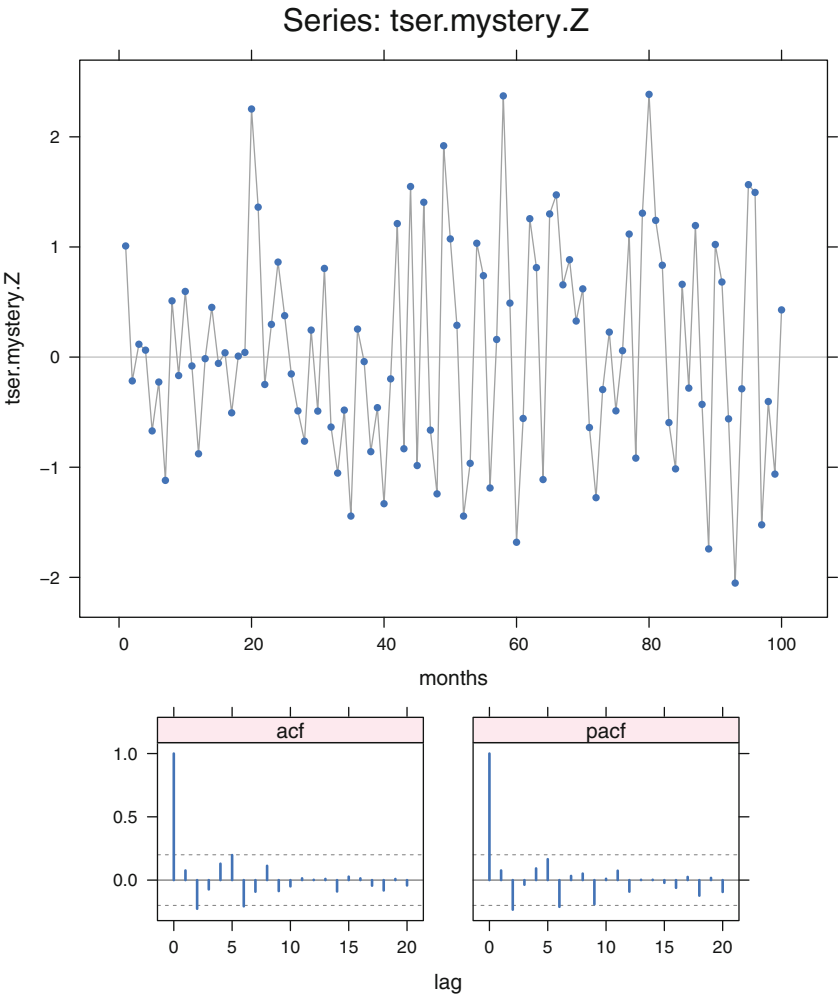

Fig. 18.15 Mystery time series Z.

18.3. Figure 18.15 shows the sequence, ACF, and PACF plots for a mystery time series Z `data(tser.mystery.Z)`. Figure 18.16 and Table 18.12 show the diagnostics and estimated coefficients obtained by fitting the 3×3 set of ARIMA$(p,0,q)$ models to the series. Table 18.13 shows the detail for the ARIMA$(1,0,1)$ model. Study the graphs and tables and explain why and how they indicate that one of these models seems better suited to explain the data than the others.

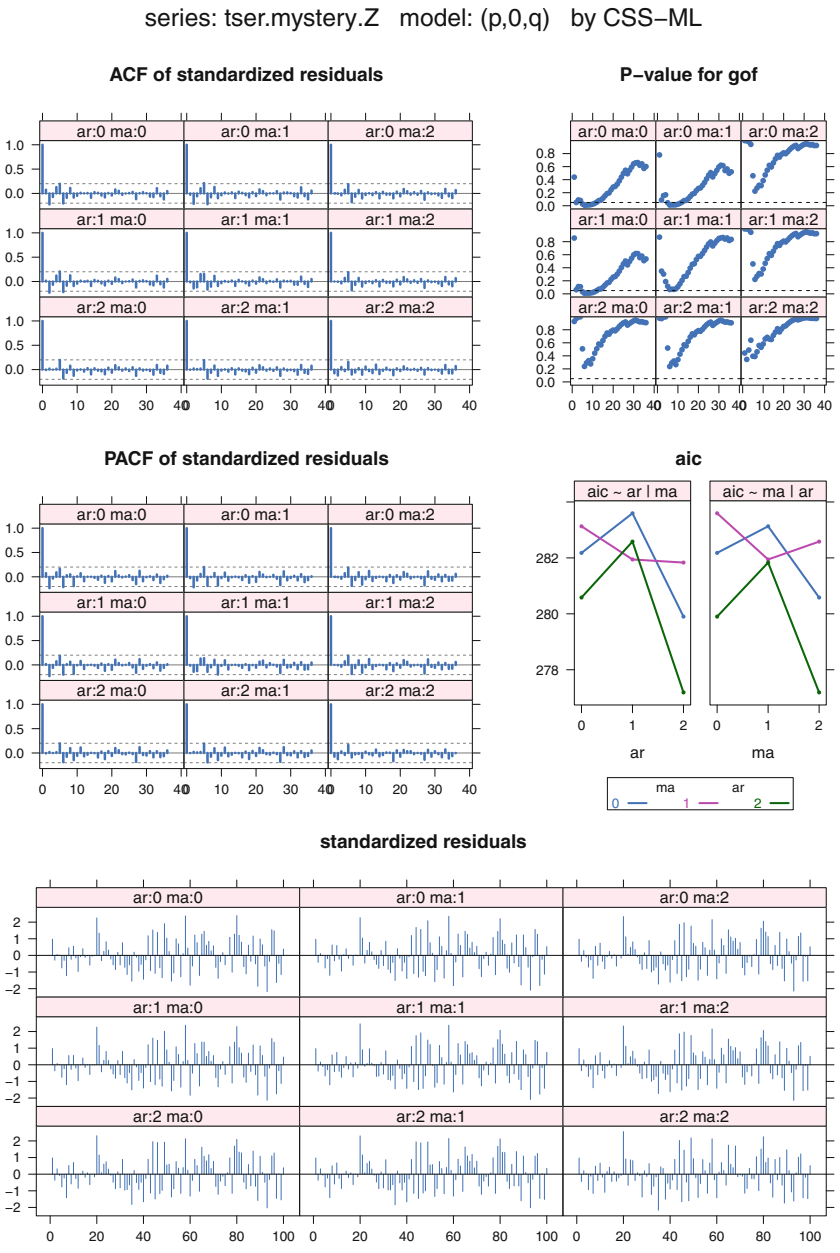

Fig. 18.16 Mystery time series Z.

Table 18.12 Mystery time series *Z*.

```
> Z.loop
$series
[1] "tser.mystery.Z"

$model
[1] "(p,0,q)"

$sigma2
       0      1      2
0 0.9454 0.9354 0.8930
1 0.9399 0.9039 0.8930
2 0.8868 0.8862 0.8088

$aic
      0     1     2
0 282.2 283.1 280.6
1 283.6 281.9 282.6
2 279.9 281.8 277.2

$coef
             ar1      ar2      ma1     ma2 intercept
(0,0,0)       NA       NA       NA      NA        NA
(1,0,0)  0.07625       NA       NA      NA   0.06379
(2,0,0)  0.09630 -0.2360       NA      NA   0.06402
(0,0,1)       NA       NA  0.13415      NA   0.06471
(1,0,1) -0.78022       NA  0.91668      NA   0.06438
(2,0,1)  0.17084 -0.2424 -0.07901      NA   0.06405
(0,0,2)       NA       NA  0.09445 -0.2177   0.06464
(1,0,2)  0.01253       NA  0.08246 -0.2183   0.06466
(2,0,2) -0.38096 -0.8567  0.57783  0.9732   0.06617

$t.coef
             ar1      ar2     ma1    ma2 intercept
(0,0,0)       NA       NA      NA     NA        NA
(1,0,0)  0.76442       NA      NA     NA    0.6083
(2,0,0)  0.98844  -2.424       NA     NA    0.7721
(0,0,1)       NA       NA  1.0571     NA    0.5905
(1,0,1) -5.17277       NA  8.8387     NA    0.6291
(2,0,1)  0.57788  -2.458 -0.2664     NA    0.7885
(0,0,2)       NA       NA  0.9615 -2.171    0.7771
(1,0,2)  0.02659       NA  0.1788 -2.139    0.7786
(2,0,2) -4.25626 -13.016  6.2170  9.171    0.6453
```

Table 18.13 Mystery time series *Z*.

```
> Z.loop[["1","1"]]

Call:
arima(x = x, order = c(1, 0, 1))

Coefficients:
         ar1    ma1  intercept
      -0.780  0.917      0.064
s.e.   0.151  0.104      0.102

sigma^2 estimated as 0.904:  log likelihood = -137,  aic = 281.9
```

18.4. The product data `data(product)` (originally from Nicholls (1979) and reproduced in Hand et al. (1994)) graphed in Figure 18.17 are the weekly sales of a plastic container used for the packaging of drugs in the United States. First differences were taken to produce the y_t, shown in Figure 18.18. The AR(1) model converged with AIC=2919, a larger number than for nonconverging models with more terms. The nonconverging models shown in Figure 18.19 and Table 18.14 showed high correlation between the estimates of the AR and MA coefficients.

a. Discuss why the above-mentioned findings and other results in Table 18.14 imply that an ARIMA$(p, 1, q)$ nonseasonal model is inappropriate for these data.

b. The peaks in the ACF and PACF plots of Figure 18.18 at 4 and 8 weeks suggest that there might be a monthly effect in this data. Examine and discuss the set of ARIMA$(p, 1, q) \times (1, 0, 0)_4$ models for these data.

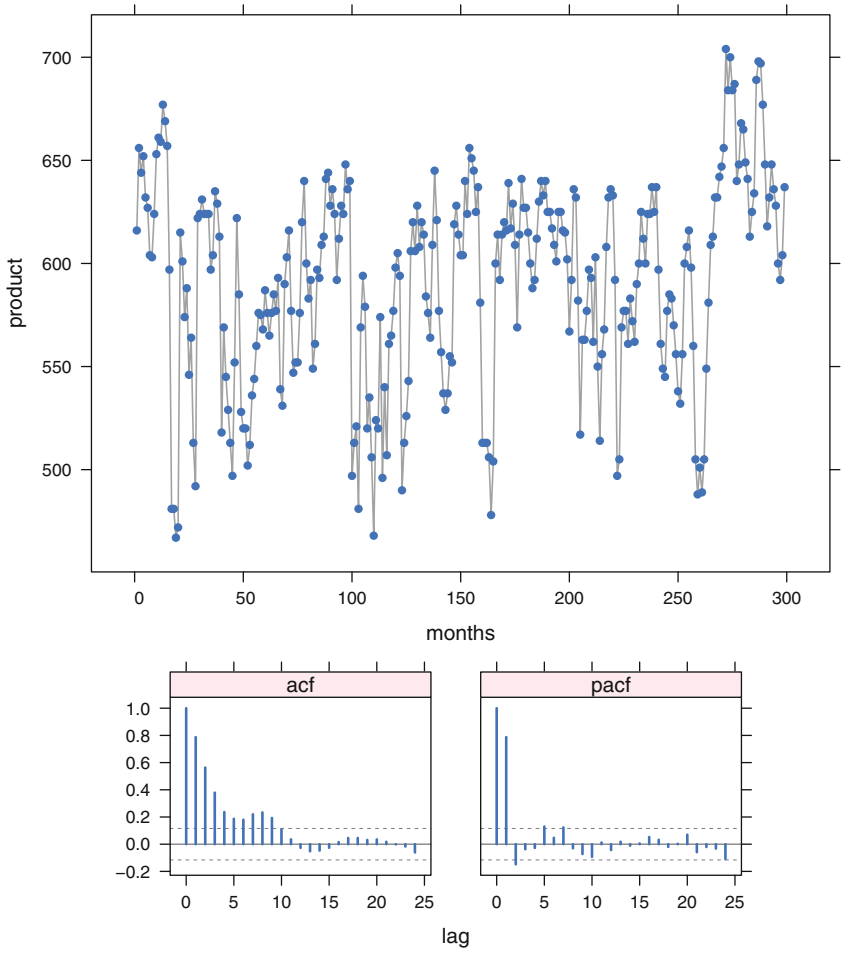

Fig. 18.17 Coordinated time series plot and ACF/PACF plots for the `product` time series: y_t. The response variable on the time series plot is weekly sales of the product.

Table 18.14 Estimation results for ARIMA($p, 1, q$) models fit to the product data.

```
> product.loop <- arma.loop(product, order=c(2,1,2))

> product.diags <- diag.arma.loop(product.loop, x=product, lag.max=60)

> product.loop
$series
[1] "product"

$model
[1] "(p,1,q)"

$sigma2
     0      1      2
0 1043 1041.8 1025.3
1 1042 1029.3  919.9
2 1033  917.8  916.3

$aic
     0    1    2
0 2919 2920 2918
1 2920 2919 2889
2 2920 2889 2890

$coef
             ar1       ar2      ma1     ma2
(0,1,0)       NA        NA       NA      NA
(1,1,0)  0.02806        NA       NA      NA
(2,1,0)  0.03066 -0.09105       NA      NA
(0,1,1)       NA        NA  0.03390      NA
(1,1,1) -0.82544        NA  0.88997      NA
(2,1,1)  0.90160 -0.15413 -0.98624      NA
(0,1,2)       NA        NA -0.02332 -0.1995
(1,1,2)  0.71744        NA -0.81250 -0.1736
(2,1,2)  1.11728 -0.32627 -1.20533  0.2183

$t.coef
            ar1    ar2      ma1     ma2
(0,1,0)      NA     NA       NA      NA
(1,1,0)  0.4833     NA       NA      NA
(2,1,0)  0.5299 -1.576       NA      NA
(0,1,1)      NA     NA   0.5341      NA
(1,1,1) -6.5556     NA   8.7970      NA
(2,1,1) 15.4832 -2.648 -61.7180      NA
(0,1,2)      NA     NA  -0.3401 -2.0824
(1,1,2) 12.9601     NA -10.9925 -2.5043
(2,1,2)  4.6121 -1.697  -4.8510  0.8844
```

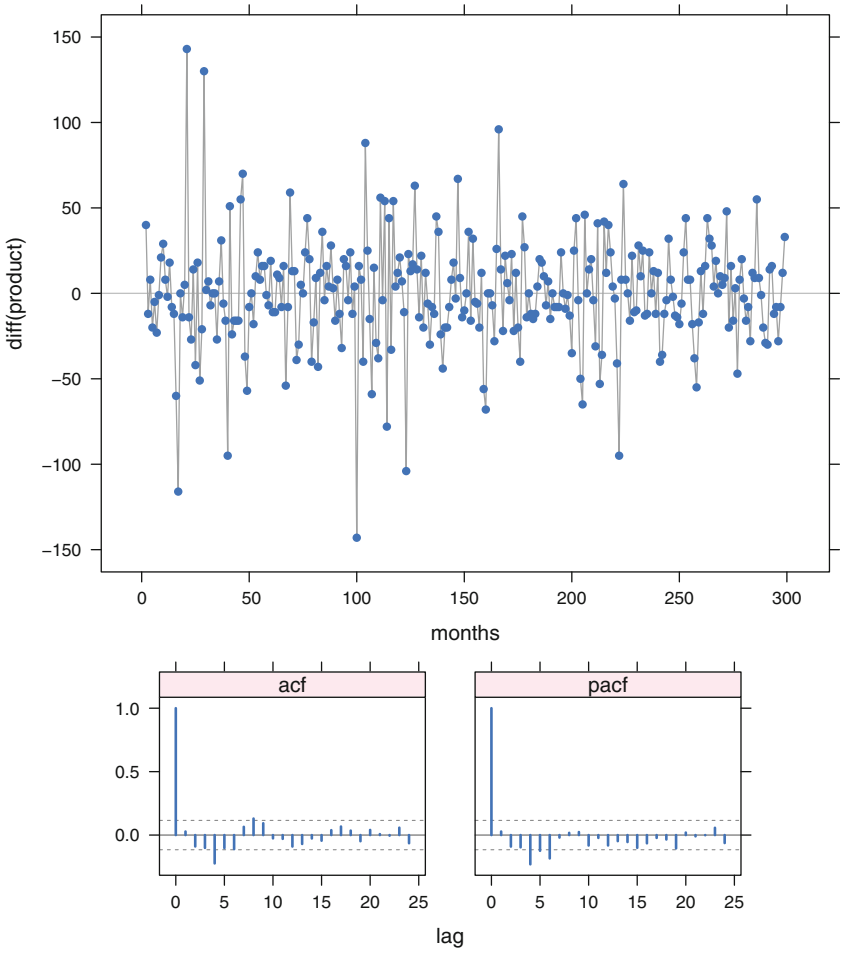

Fig. 18.18 Coordinated time series and ACF/PACF plots for the differenced `product` time series: ∇y_t.

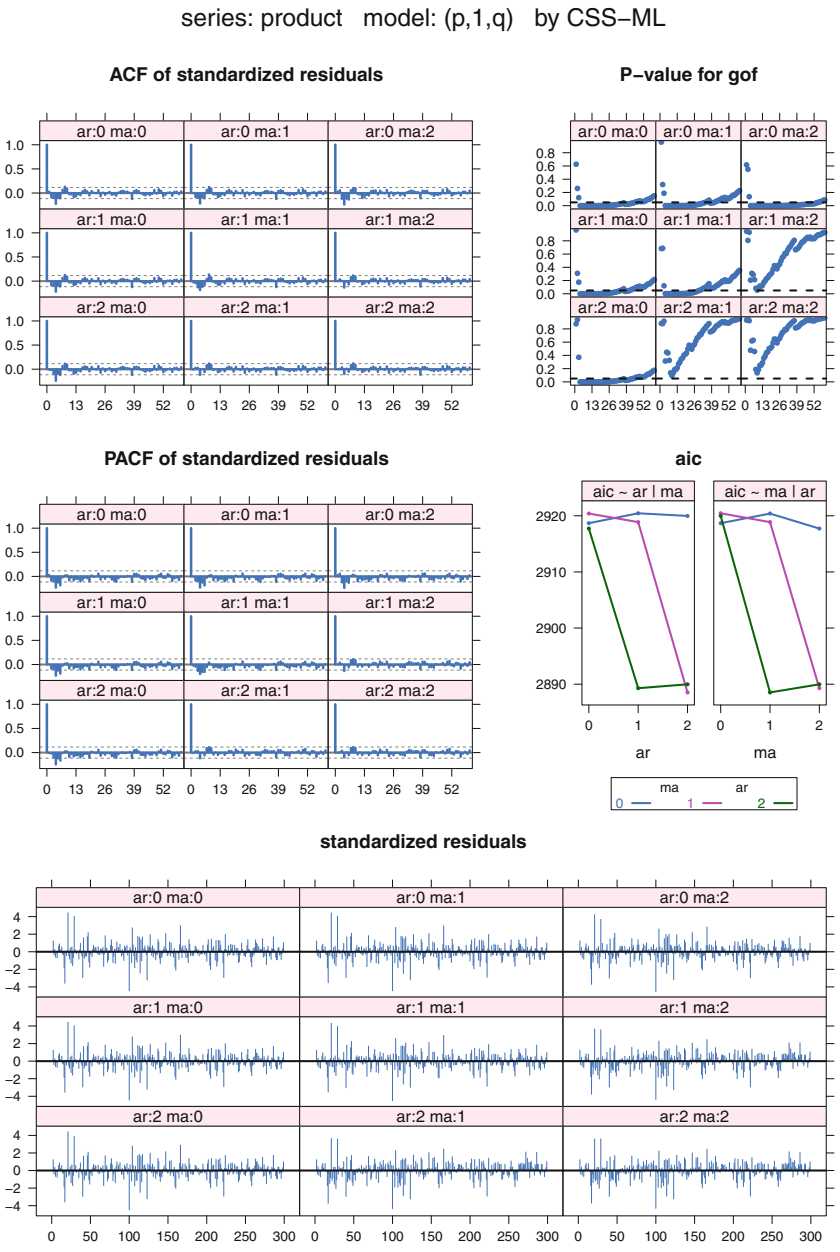

Fig. 18.19 Diagnostic plots for the set of models ARIMA(p,1,q) fit to the `Product` data by maximum likelihood. Each set of nine panels is systematically structured in a 3×3 array with rows indexed by the number of AR parameters p and columns by the number of MA parameters q. All nine panels in a set are scaled identically. The AIC has been plotted as a pair of interaction plots: AIC plotted against q using line types defined by p; and AIC plotted against p, using line types defined by q.

18.5. Figures 18.20, 18.21, and 18.22 and Tables 18.15 and 18.16 show the mean monthly air temperature in degrees Fahrenheit from January 1920 to December 1939 at Nottingham. R users can use the `nottem` data in the pkgdatasets package. We first got the data from Venables and Ripley (1997). The original source is "Meteorology of Nottingham" in *City Engineer and Surveyor*. We show the original series, the seasonally differenced series, the diagnostic display from the series of models ARIMA$(p, 0, q) \times (2, 1, 0)_{12}$, and numerical results from the set of all nine models table and detail on the recommended model ARIMA$(1, 0, 0) \times (2, 1, 0)_{12}$.

a. What are the most evident features of the plot of the original data?

b. Compare the plot of the seasonally differenced data to the original plot. What structure was captured by the differencing? What remains?

c. Compare the recommended model ARIMA$(1, 0, 0) \times (2, 1, 0)_{12}$ to the next most likely model ARIMA$(2, 0, 0) \times (2, 1, 0)_{12}$. Do you agree that the `ar(2)` term is not needed? Why?

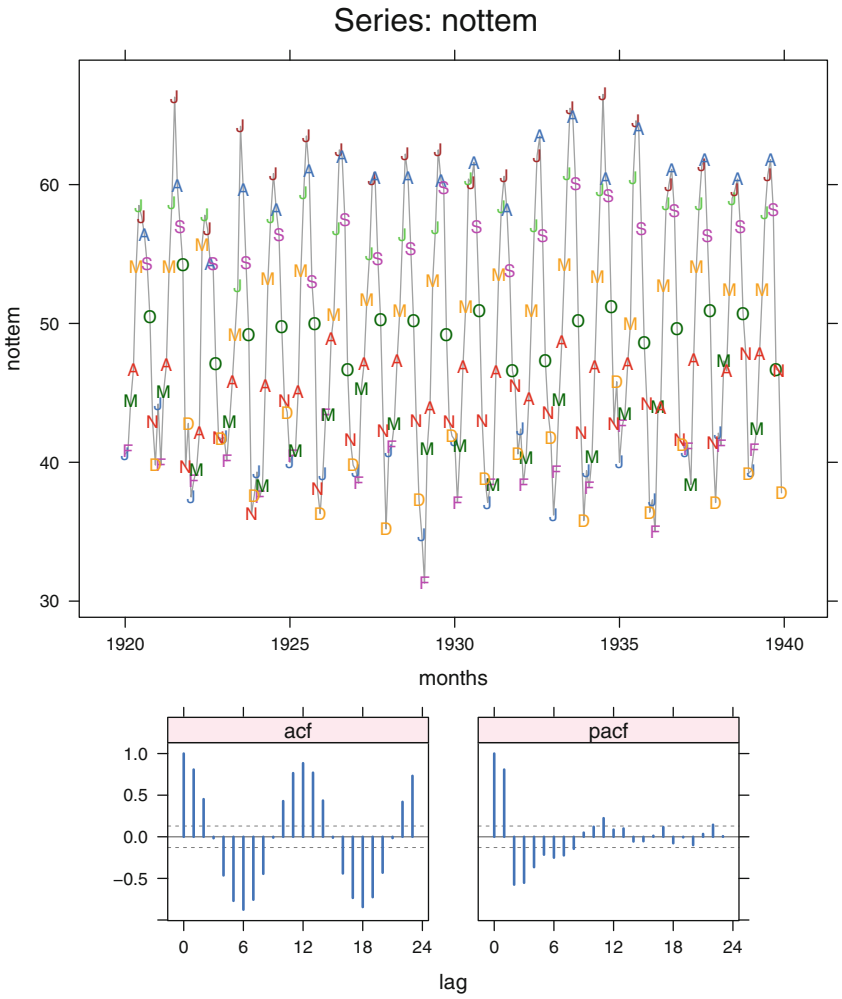

Fig. 18.20 Mean monthly air temperature in degrees Fahrenheit from January 1920 to December 1939 at Nottingham.

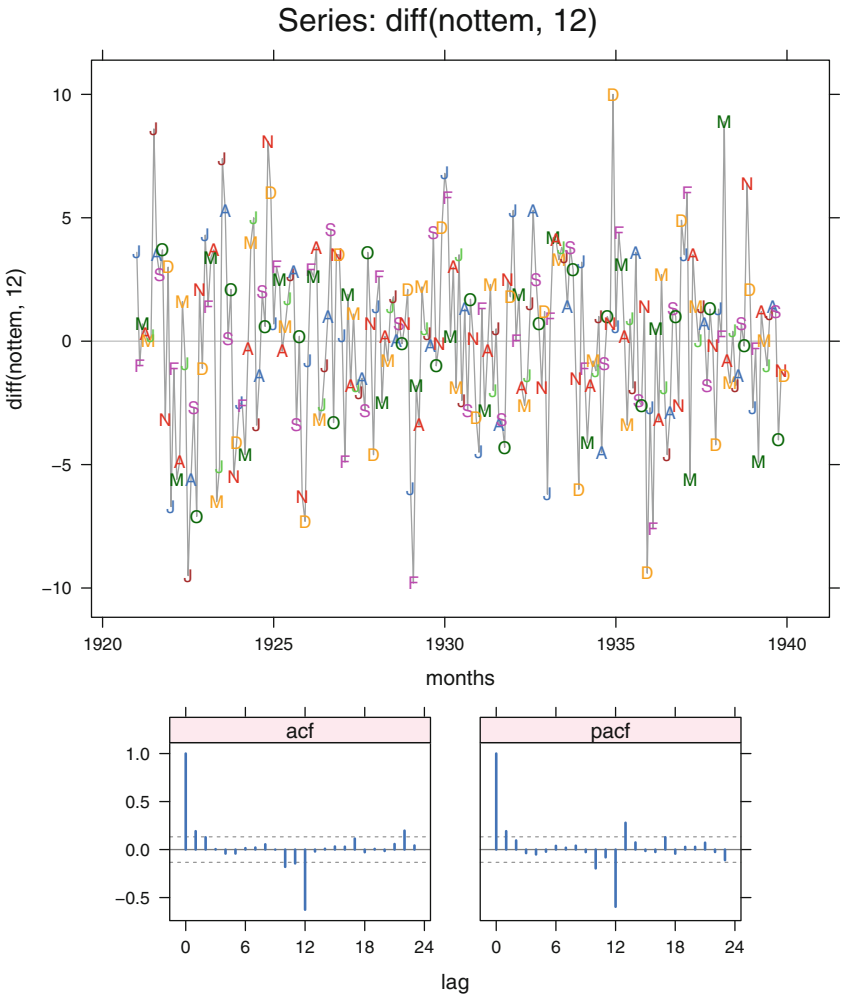

Fig. 18.21 Seasonal differences of mean monthly air temperature in degrees Fahrenheit from January 1920 to December 1939 at Nottingham.

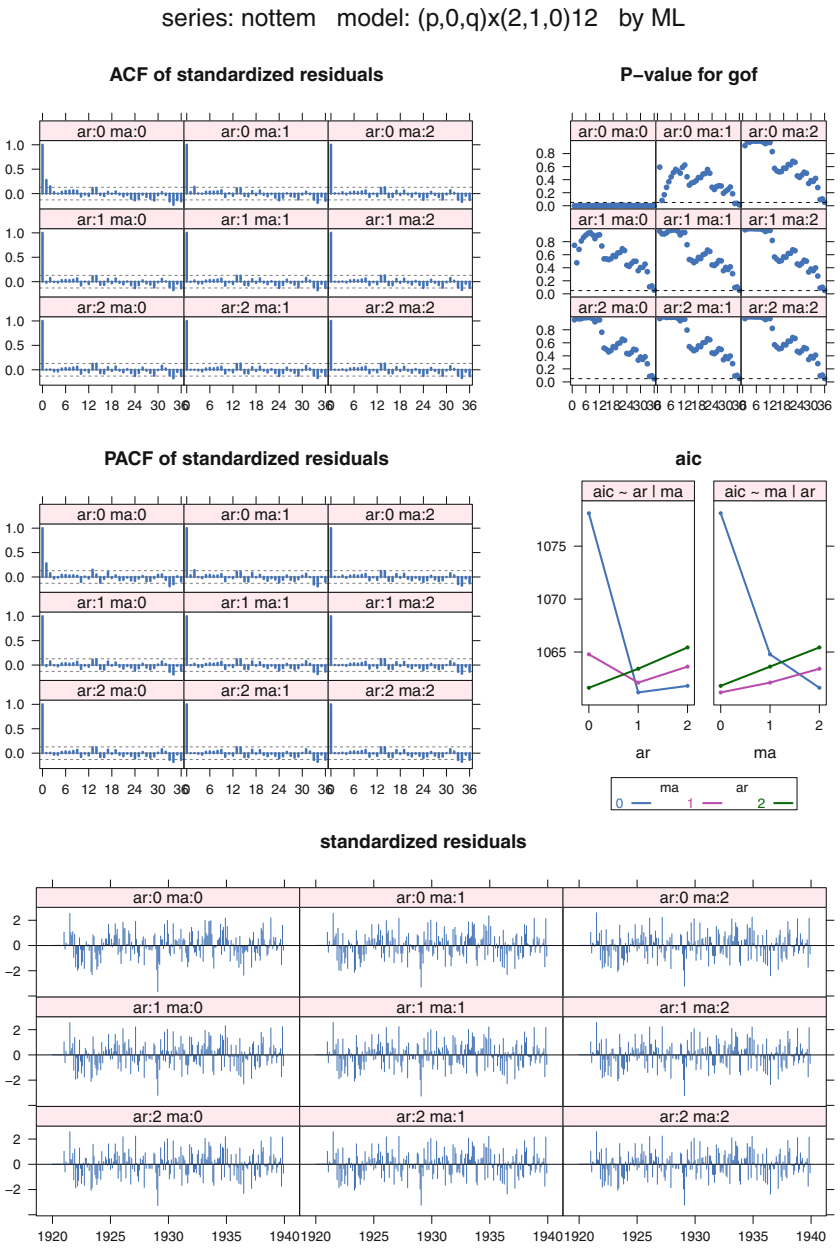

Fig. 18.22 Mean monthly air temperature in degrees Fahrenheit from January 1920 to December 1939 at Nottingham Castle.

Table 18.15 Nottingham temperature—models ARIMA$(p, 0, q) \times (2, 1, 0)_{12}$.

```
> nottem.loop <- arma.loop(nottem,  order=c(2,0,2),
+                          seasonal=list(order=c(2,1,0), period=12),
+                          method="ML")

> print(nottem.loop, digits=4)
$series
[1] "nottem"

$model
[1] "(p,0,q)x(2,1,0)12"

$sigma2
      0     1     2
0 6.219 5.799 5.661
1 5.702 5.674 5.656
2 5.666 5.661 5.656

$aic
     0    1    2
0 1078 1065 1062
1 1061 1062 1063
2 1062 1064 1065

$coef
                       ar1     ar2      ma1    ma2    sar1    sar2
(0,0,0)x(2,1,0)12       NA      NA       NA     NA -0.8220 -0.2931
(1,0,0)x(2,1,0)12 0.285599      NA       NA     NA -0.8598 -0.2963
(2,0,0)x(2,1,0)12 0.261443 0.07937       NA     NA -0.8602 -0.3074
(0,0,1)x(2,1,0)12       NA      NA  0.23606     NA -0.8505 -0.2866
(1,0,1)x(2,1,0)12 0.478197      NA -0.20979     NA -0.8603 -0.3033
(2,0,1)x(2,1,0)12 0.001805 0.15577  0.26037     NA -0.8608 -0.3100
(0,0,2)x(2,1,0)12       NA      NA  0.25718 0.1575 -0.8607 -0.3131
(1,0,2)x(2,1,0)12 0.169231      NA  0.09343 0.1200 -0.8611 -0.3127
(2,0,2)x(2,1,0)12 0.118763 0.01077  0.14252 0.1212 -0.8614 -0.3153

$t.coef
                       ar1    ar2     ma1   ma2   sar1    sar2
(0,0,0)x(2,1,0)12       NA     NA      NA    NA -13.07  -4.418
(1,0,0)x(2,1,0)12 4.451829     NA      NA    NA -13.46  -4.443
(2,0,0)x(2,1,0)12 3.899219 1.1772      NA    NA -13.54  -4.596
(0,0,1)x(2,1,0)12       NA     NA  4.0370    NA -13.32  -4.275
(1,0,1)x(2,1,0)12 2.668022     NA -1.0576    NA -13.52  -4.556
(2,0,1)x(2,1,0)12 0.003237 0.9644  0.4644    NA -13.56  -4.620
(0,0,2)x(2,1,0)12       NA     NA  3.9101 2.302 -13.58  -4.661
(1,0,2)x(2,1,0)12 0.448103     NA  0.2501 1.008 -13.58  -4.650
(2,0,2)x(2,1,0)12 2.887888    NaN  1.0990 1.015 -26.55 -29.916


> nottem.diag <-
+    rearrange.diag.arma.loop(diag.arma.loop(nottem.loop, nottem))
```

Table 18.16 Nottingham temperature—recommended model ARIMA$(1,0,0)\times(2,1,0)_{12}$.

```
> nottem.loop[["1","0"]]

Call:
arima(x = x, order = c(1, 0, 0), seasonal = list(order = c(2, 1, 0),
      period = 12),  method = "ML")

Coefficients:
        ar1    sar1    sar2
      0.286  -0.860  -0.296
s.e.  0.064   0.064   0.067

sigma^2 estimated as 5.7:  log likelihood = -526.6,  aic = 1061
```

18.6. We have a time series of size $n = 100$ for which we have determined that we have an ARIMA(1,0,0) model and have estimated $\hat{\mu} = 15$, $\hat{\phi} = .2$, and $\hat{\sigma}^2 = 3$. The last few observations in the series are

t	97	98	99	100
X_t	13	15	18	17

Forecast, with 95% forecast intervals, the values $\hat{X}_{101}$ and $\hat{X}_{102}$.

18.7. We have a nonseasonal time series in the dataset `data(tsq)` covering 100 periods. The time series and its ACF and PACF plots are displayed in Figure 18.23. Table 18.17 contains the R output from a 3×3 set of ARIMA models fit to the data. The `tsdiagplot` for these data is in Figure 18.24. Use this information to answer the following questions:

a. Recommend the $(p,0,q)$ order for an ARIMA modeling of these data.

b. Write out the equation for the best-fitting model following your recommendation in part (a).

c. Use your model and the coefficient information in the R output to produce forecasts and 95% forecast intervals for the value of this series in periods 101 and 102.

Table 18.17 Three by three set of ARIMA models for Exercise 18.7.

```
> tsq.loop <- arma.loop(tsq, order=c(2,0,2))

> tsq.loop
$series
[1] "tsq"

$model
[1] "(p,0,q)"

$sigma2
      0      1      2
0 1.330 0.9469 0.9267
1 1.173 0.9328 0.8730
2 1.003 0.9175 0.8715

$aic
      0     1     2
0 316.3 285.2 285.0
1 305.8 285.6 283.2
2 292.6 286.0 284.9

$coef
            ar1      ar2     ma1      ma2 intercept
(0,0,0)      NA       NA      NA       NA        NA
(1,0,0)  0.3444       NA      NA       NA   0.09045
(2,0,0)  0.4894 -0.38808      NA       NA   0.08309
(0,0,1)      NA       NA  0.7524       NA   0.08843
(1,0,1) -0.1523       NA  0.8031       NA   0.09029
(2,0,1) -0.1087 -0.15017  0.7447       NA   0.08999
(0,0,2)      NA       NA  0.6043 -0.1653   0.09055
(1,0,2)  0.8468       NA -0.2290 -0.7710   0.12449
(2,0,2)  0.7879  0.06726 -0.2076 -0.7924   0.12323

$t.coef
            ar1     ar2    ma1     ma2 intercept
(0,0,0)      NA      NA     NA      NA        NA
(1,0,0)  3.6551      NA     NA      NA    0.5505
(2,0,0)  5.1886 -4.0813     NA      NA    0.7431
(0,0,1)      NA      NA 11.139      NA    0.5208
(1,0,1) -1.2440      NA 12.664      NA    0.5992
(2,0,1) -0.8042 -1.2726  7.520      NA    0.6785
(0,0,2)      NA      NA  5.272  -1.493    0.6549
(1,0,2) 14.4918      NA -3.308 -11.612    3.9454
(2,0,2)  6.4214  0.5445 -2.841 -11.185    3.7167
```

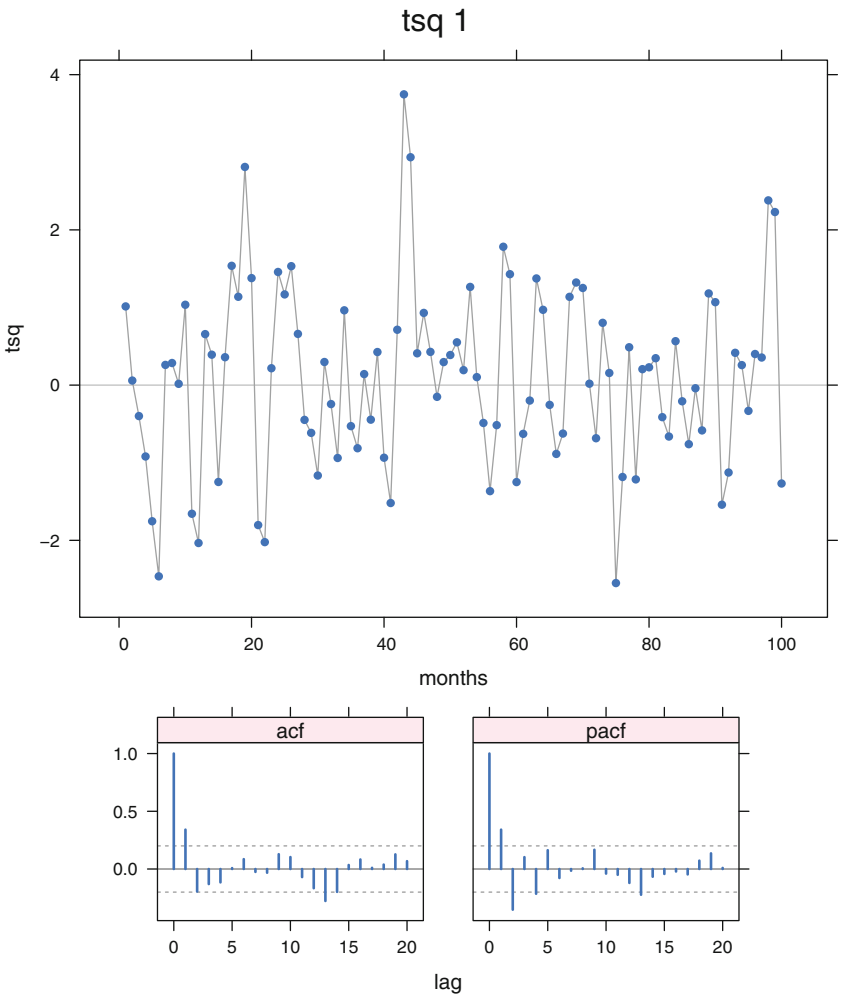

Fig. 18.23 Time series and its ACF and PACF plots for Exercise 18.7.

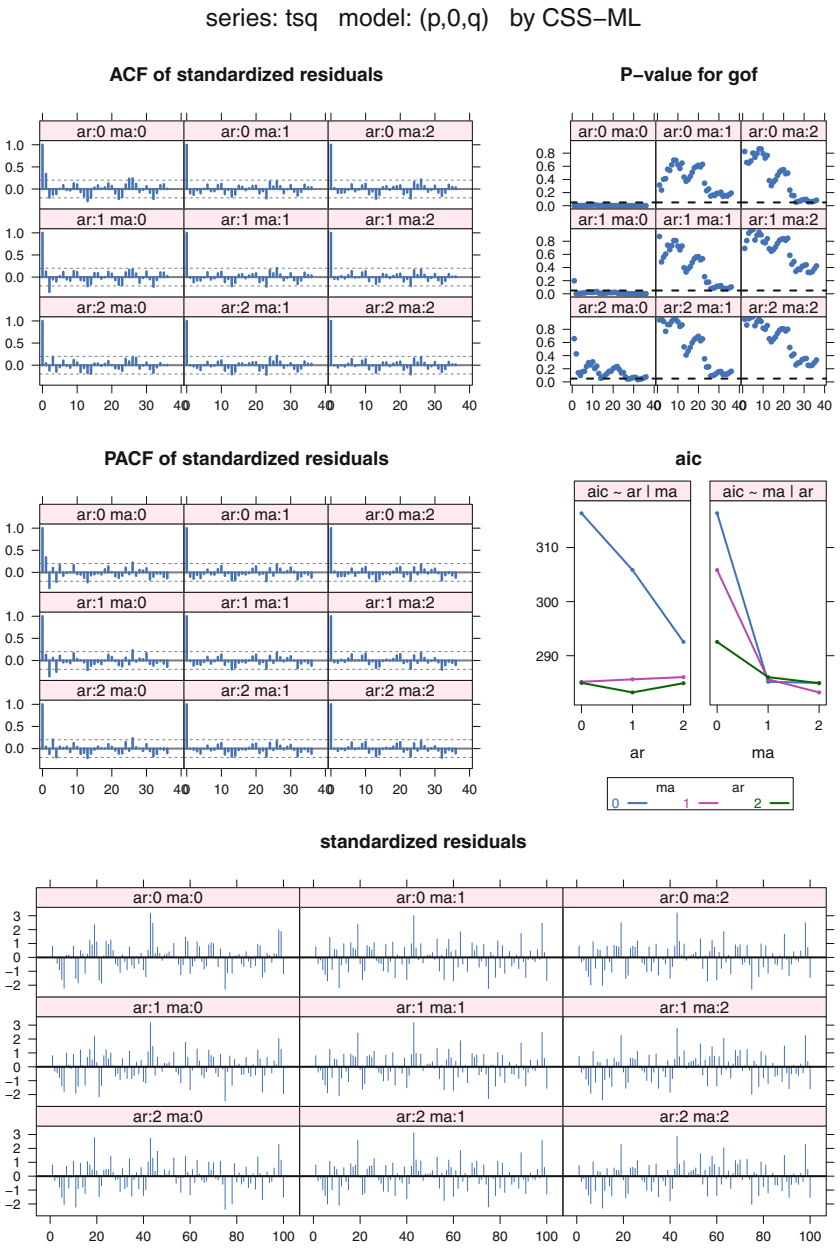

Fig. 18.24 Diagnostic plots for the 3×3 set of ARIMA models in Exercise 18.7.

18.8. Figure 18.25 contains time series, ACF, and PACF plots for monthly data on the thickness of the ozone layer (measured in Dobson units) at Arosa, Switzerland, from September 1931 through November 1953. Note the labeling of the months of the year (J = January, June, or July, F = February, etc.) at the plot points. The data in the dataset `data(ozone)` are from Andrews and Herzberg (1985), Table 12.1.

a. Comment on the seasonal nature of the time series plot and discuss how this is consistent with what you see in the ACF plot.

b. Notice that the variability of the series appears to increase, at least temporarily, in the early 1940s and around 1952 to 1953. For each of these periods, *identify historical events* that potentially impacted on the atmosphere to produce this increased variability.

18.9. $n = 100$ and $\bar{Z} = 25$. See Figure 18.26.

a. Identify a tentative underlying model in explicit form and justify your model.

b. Propose possible preliminary parameter estimates for your model.

c. Assume the residual sum of squares from the fitting of your model is 256, and $Z_{98} = 24, Z_{99} = 26, Z_{100} = 25$. Compute your forecasts for Z_{101} and Z_{102} and their 95% forecast intervals.

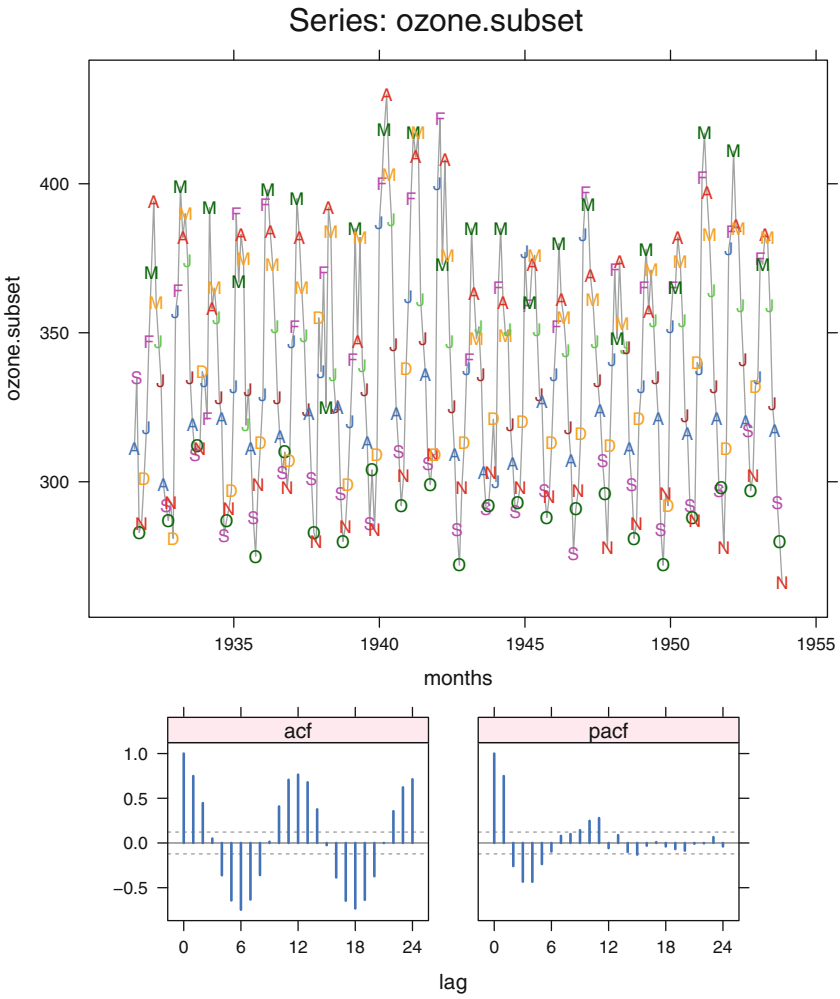

Fig. 18.25 Thickness of the ozone layer (measured in Dobson units) at Arosa, Switzerland, from September 1931 through November 1953.

	lag									
	1	2	3	4	5	6	7	8	9	10
acf(Z_t)	0.80	0.61	0.47	0.40	0.31	0.21	0.18	0.11	0.06	0.01
pacf(Z_t)	0.80	0.08	0.00	−0.11	0.00	−0.12	0.07	0.05	0.01	0.02

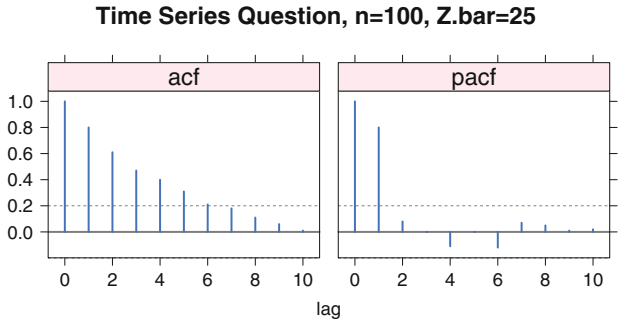

Fig. 18.26 ACF and PACF for Exercise 18.9, $n = 100$ and $\bar{Z} = 25$. The same information is presented in both tabular and graphical form.

	lag									
	1	2	3	4	5	6	7	8	9	10
acf(Z_t)	0.93	0.92	0.90	0.90	0.87	0.86	0.85	0.84	0.82	0.80
acf(∇Z_t)	−0.57	−0.10	0.12	0.06	−0.12	0.09	0.05	−0.01	0.02	0.03

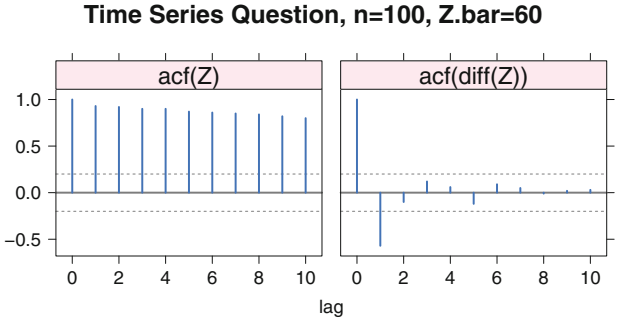

Fig. 18.27 ACF and PACF for Exercise 18.10, $n = 100$ and $\bar{Z} = 60$. The same information is presented in both tabular and graphical form.

18.10. $n = 100$ and $\bar{Z} = 60$. Identify a tentative underlying model in explicit form and justify your model. See Figure 18.27.

18.11. $n = 100$ and $\bar{Z} = 55$. Identify a tentative underlying model in explicit form and justify your model. See Figure 18.28.

18.12. Time series data differs from any other data type we have discussed in one important characteristic: The observations are not independent. What are the implications of that difference for modeling time series data? Be sure to discuss implications for each of

a. Modeling

b. Estimation

c. Prediction

	lag									
	1	2	3	4	5	6	7	8	9	10
acf(Z_t)	0.99	0.94	0.87	0.81	0.75	0.65	0.55	0.53	0.43	0.40
acf(∇Z_t)	0.43	0.28	0.51	0.80	0.65	0.44	0.31	0.77	0.30	0.20
acf($\nabla_4 Z_t$)	0.72	0.67	0.55	0.32	0.38	0.23	0.24	0.23	0.18	0.13
acf($\nabla\nabla_4 Z_t$)	0.30	0.07	0.32	0.50	0.20	0.01	−0.05	−0.01	0.02	0.03

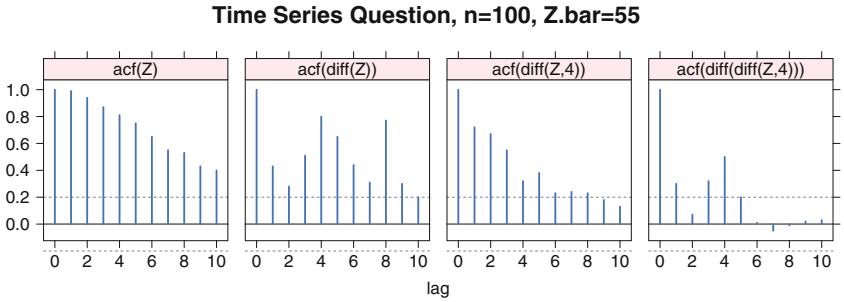

Fig. 18.28 ACF and PACF for Exercise 18.11, $n = 100$ and $\bar{Z} = 55$. The same information is presented in both tabular and graphical form.

18.13. Figure 18.29 shows the "United States of America Monthly Employment Figures for Males Aged 16–19 Years from 1948 to 1981". Dataset `data(employM16)` is Table T.65.1 from Andrews and Herzberg (1985). What are the features of this plot that you would try to capture in a time series model? Comment on

a. Seasonality

b. Trend

c. Aberrations

18.A Appendix: Construction of Time Series Graphs

This section discusses the technical aspects of the construction of the set of plots used to check the validity of the proposed model. The interpretation of the plots, and the discussion of how to use them to help identify the model that best fits the data, appear in Sections 18.6 and 18.8.

The graphical display techniques demonstrated in Sections 18.6 and 18.8 were developed by Heiberger and Teles (2002). The R functions from the **HH** package used to produce these displays are described in help files `?tsacfplots` and `?tsdiagplot`.

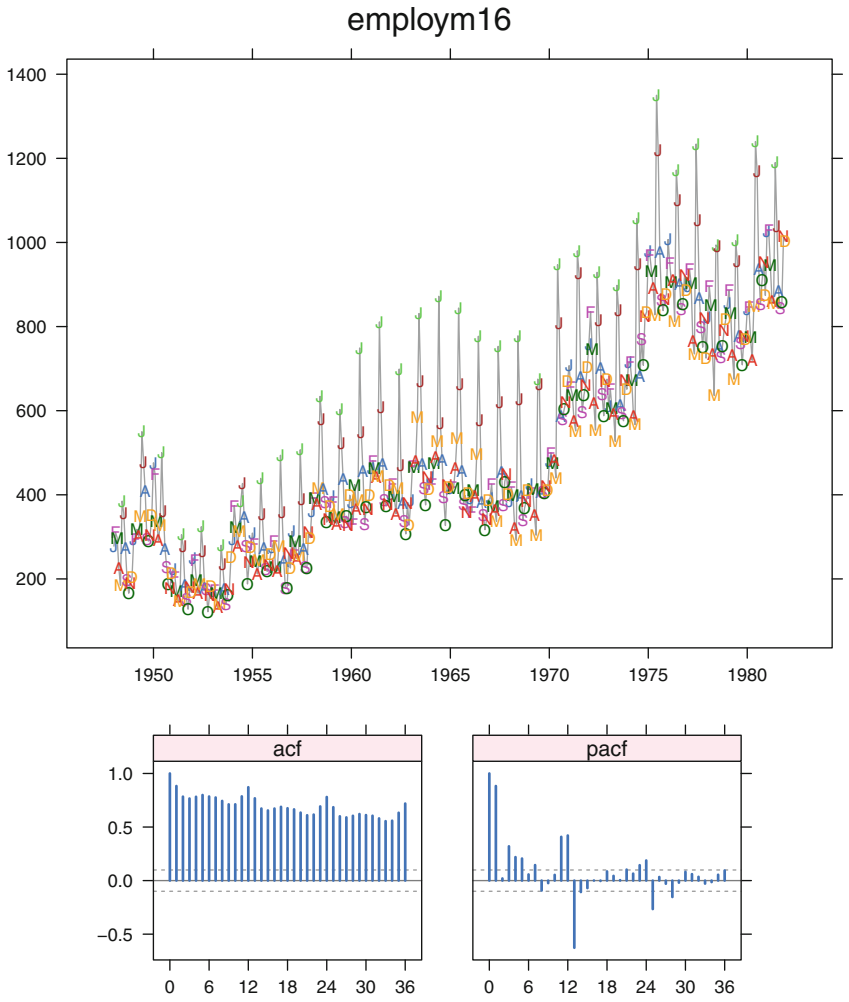

Fig. 18.29 United States of America Monthly Employment Figures for Males Aged 16–19 Years from 1948 to 1981, for use with Exercise 18.13.

The set of plots in Figure 18.8 consists of the residual ACF and PACF, the portmanteau goodness-of-fit test statistic (GOF), the standardized residuals, and the Akaike information criterion (AIC). The panels in the first four sets of plots are indexed by the number of nonseasonal ARMA parameters p and q for fixed values of the seasonal parameters P and Q. The AIC plot uses p and q as plotting variables. The orders of differencing and the orders of the autoregressive and moving average operators (both seasonal and nonseasonal) have been limited to $0 \le p, d, q, P, D, Q \le 2$. While this limitation is usually reasonable in practice, it is not inherent in the software.

Each set of nine panels is systematically structured in a 3×3 array indexed by the number of AR parameters and MA parameters. All nine panels in a set are scaled identically. Thus the reader can scan a row or column of the array of panels and see the effect of adding one more parameter to either the AR or MA side of the model.

Traditionally (that is, as constructed by the standard R `tsdiag`), the plots coordinated in Figure 18.8 are shown on nine separate pages, one page for each model. The standard display shows the standardized residuals, the residual ACF and PACF plots, and the portmanteau goodness-of-fit test. The nine sets of plots, each associated with a different model, will not necessarily be scaled alike. Even the GOF and ACF/PACF plots for the same model may have different lag scales.

Labeling the axis in months and putting the residual ACF and PACF plots and the GOF plot on the same set of lags make it easy to compare the plots for different models. In this example it is easy to see that something is happening at `lag=12` months. The AIC plots for all the models in Figure 18.8 are similar, with AIC≈315. The AIC has been plotted as a pair of interaction plots: AIC plotted against q, the number of nonseasonal MA parameters, using line types defined by p, the number of nonseasonal AR parameters; and AIC plotted against p, using line types defined by q. These plots enable us to study the magnitudes of the differences in AIC of competing models.

18.A.1 Characteristics of This Presentation of the Time Series Plot

- Individual points are identified with a letter indicating the position of each observation according to the frequency of collection of the data. The user can control the choice of plotting characters. The default characters used are dependent on the frequency of collection of the data. For example, when the frequency is 12, the default plotting characters are the month abbreviations `J,F,M,A,M,J,J,A,S,O,N,D`. Otherwise they are chosen from the beginning of the lower case alphabet `letters()`.
- The plotting characters are an explicit argument and can be chosen by the user (with `pch.seq`), or suppressed entirely with `type="l"`.
- Color is often very helpful with the time series plots. Color plots show the seasonal pattern more strongly than the black and white. Figure 18.4 has a clear pattern of gold-May and red-April along its top and green-October along the bottom. The first differences in Figure 18.6 show a clear pattern of blue-August along the bottom and color-coded June–July–August–September below the axis. Figure 18.7 shows random behavior in colors.

18.A.2 Characteristics of This Presentation of the Sample ACF and PACF Plots

- The axes are coordinated and have the same scale.
- Lags are indicated in appropriate units (for example, months for monthly series).
- The ACF and PACF plots consistently both show, or do not show (at the user's option), the spike for correlation=1 at `lag=0`.
- The default tick marks are related to the frequency of collection of the data. The user has control over tick mark location.
- Most of the plotting surface is occupied by the body of the plot, and the amount of surface used for labeling is minimized.

We point out that the individual plots are accessible to the user. They can be placed on their own pages or displayed with other relative spacings. For details, see the function `HH:::print.tsacfplots`.

18.A.3 Construction of Graphical Displays

This section shows how to construct the two display types presented in this chapter. For brevity, only Figures 18.7 and 18.9 are described.

Figure 18.7, a single display with subgraphs, is constructed with the single command:

```
tsacfplots(diff(diff(co2,1), 12))
```

The figure uses the majority of the plotting surface to display the time series itself and a minority of the plotting surface to display the ACF and PACF plots drawn to the same scale.

All models in the family of ARIMA models under investigation are fit with a single command specified in standard R time series model notation:

```
ddco2.loopPQ <-
     arma.loop(co2,
               order=c(2,1,2),
               seasonal=list(order=c(0,1,1), period=12))
```

Figure 18.9 plots the family of models, again as a single display with coordinated sets of subgraphs, with another single command:

```
tsdiagplot(armas=co2.loop)
```

The series of plots in each set of subgraphs is displayed in the same systematic order. All plots of the same form are displayed to the same scale.

Fine control of plotting options and labeling is possible with optional arguments to the `tsacfplots` and `tsdiagplot` functions. Each of the individual subgraphs is also directly accessible to the user.

Formal and systematic display of a series of models makes it easy to recognize the structural differences in the series of models and to compare them.

18.A.4 Functions in the HH package for R

Several functions are provided and described here in terms of their role in the modeling. In addition to these functions, there are unexported functions that the primary functions call to do much of the work.

Primary Functions in the **HH** Package

`tsacfplots`: Provides a single display (of the form of Figure 18.7) with the times series plot central and both the ACF and PACF plots on the same scale. It does so by calling `seqplot` (equivalent to `ts.plot` but with much finer control of labeling options) for the time series plot and then `acf.pacf.plot` for the coordinated ACF and PACF plots. These in turn are constructed by the R routine `acf`.

`arma.loop`: Takes a time series and a model statement of the form

$$(p_{\max}, d, q_{\max}) \times (P, D, Q)_{\texttt{period}}$$

It then loops through the family of models indexed by the model parameters $1{:}p_{\max}$ and $1{:}q_{\max}$, with d, P, D, Q held constant. Results are stored in a list indexed by the values of p and q.

`arma.loop` also permits the model statement (note that order matters)

$$(P_{\max}, D, Q_{\max})_{\texttt{period}} \times (p, d, q)$$

and then loops through the family of models indexed by the model parameters $1{:}P_{\max}$ and $1{:}Q_{\max}$, with p, d, q, D held constant. Results are stored in a list indexed by the values of P and Q.

`diag.arma.loop`: Produces an indexed list of the `arima.diag` results for each model in the result of the `arma.loop`. Diagnostics are calculated on the boundary values of the parameters p and q, and in particular for those functions defined in the special case $(p, d, q) = (0, 0, 0)$.

`tsdiagplot`: Takes a time series and a model statement and calls all the diagnostic plot routines. It makes sensible default choices for all the arguments and produces a graph similar to Figure 18.9. For printouts of any of the numerical tables, or finer control over the layout and labeling of the plots, the user should study the more detailed illustrations of function use in the `demo("tsamstat")`.

Print Methods

`print.arma.loop` and `summary.arma.loop`: Produce tables similar to Table 18.5 from the result of the `arma.loop` function.

Individual Plot Functions

Each of the subgraphs in `tsacfplots` and `tsdiagplot` is directly accessible to the users. Each is fully parameterized.

`tsacfplots`: Figures 18.4 and 18.7.

`seqplot`	Time series
`acf.pacf.plot`	Coordinated ACF and PACF

`tsdiagplot`: Figures 18.8 and 18.9.

`acfplot`	ACF and PACF of residuals
`residplot`	Standardized residuals
`gofplot`	Portmanteau goodness-of-fit statistic (GOF)
`aicsigplot`	Interaction plot of AIC or σ^2

`seqplotForecast`: Figure 18.10. Data, forecasts, and confidence bands.

Additional functions

Not for direct use by users.

`rearrange.diag.arma.loop`: Rearranges the list of diagnostics indexed by model into a list of matrices of diagnostics, each matrix indexed by the models. The sole purpose of this rearrangement is for plotting.