Lecture 12

ARMA Case Study & Autoregressive Integrated Moving Average (ARIMA) Models

Reading: Cryer and Chen (2008): Chapter 5.1-5.3

MATH 4070: Regression and Time-Series Analysis

Whitney Huang
Clemson University

Agenda

1. **ARMA Case Study**

2. **ARMIA**

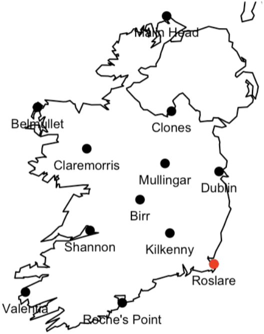
A Modeling Case Study of Ireland Wind Data

(Courtesy of Peter Craigmile's time series lecture notes)

Data Description [Haslett & Raftery, 1989 [1]]

Twelve wind stations collected daily readings over 18 years (from 1961 to 1978). Wind speeds were measured in knots (1 knot = 0.5148 $\frac{m}{s}$)

We will focus on the wind data from 1965-1969 at the Rosslare station

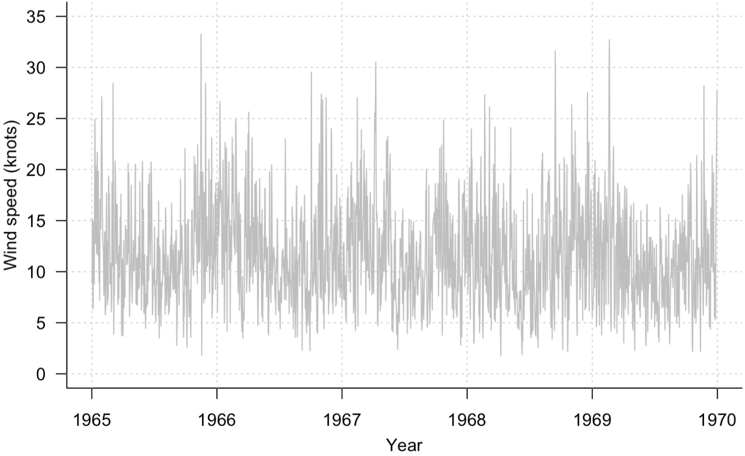


Modeling procedure:

- Exploratory analysis
- Model and remove the trend and seasonal components
- ARMA model identification, fitting, and selection
- Perform forecast

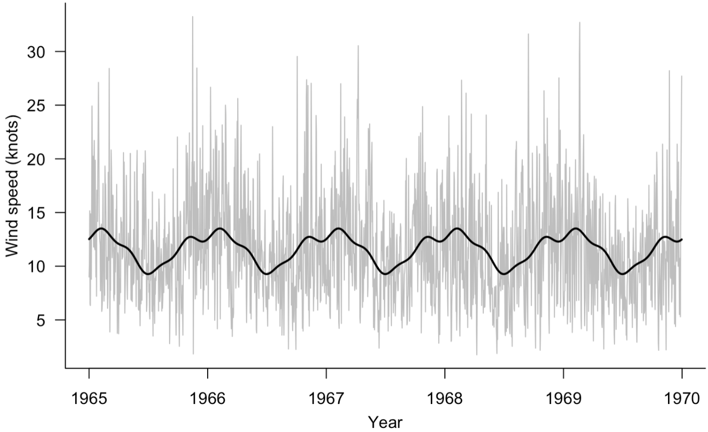
[1] Haslett, J., & Raftery, A. E. (1989). Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. Journal of the Royal Statistical Society: Series C, 38(1), 1-21.

Wind Speed Time Series at Rosslare Station



- No clear trend
- Seasonal Pattern

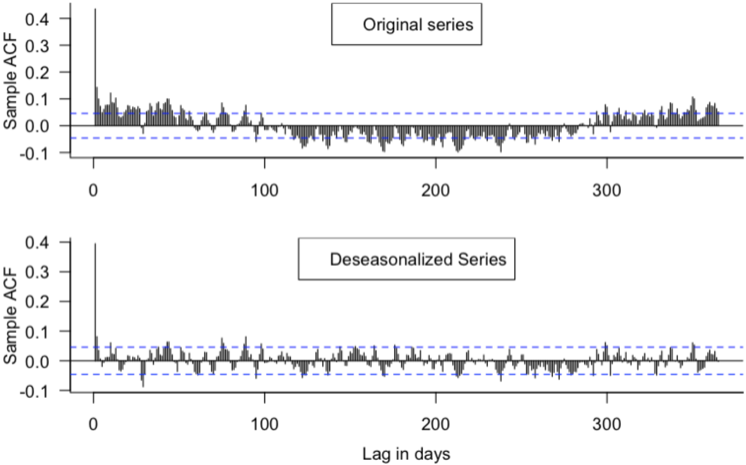
Estimating the Season Pattern



Here we use harmonic regression with 4 harmonics per year to model the seasonal components

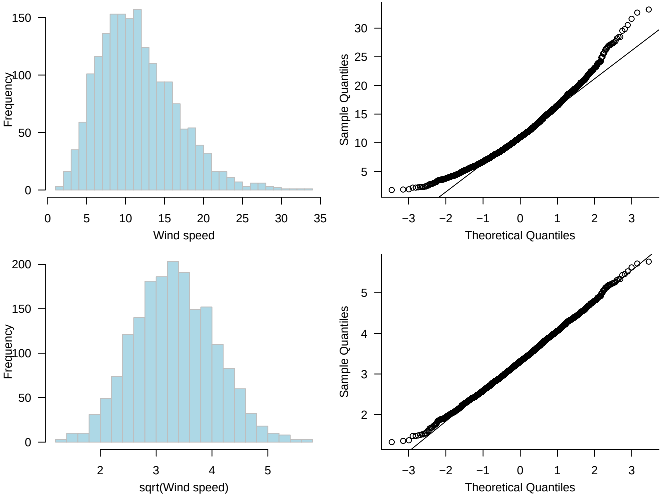
$$s_t = \beta_0 + \sum_{j=1}^{4} \left(\beta_{1j} \cos(2\pi jt) + \beta_{2j} \sin(2\pi jt) \right)$$

ACF Plots: Original and Deseasonalized Series



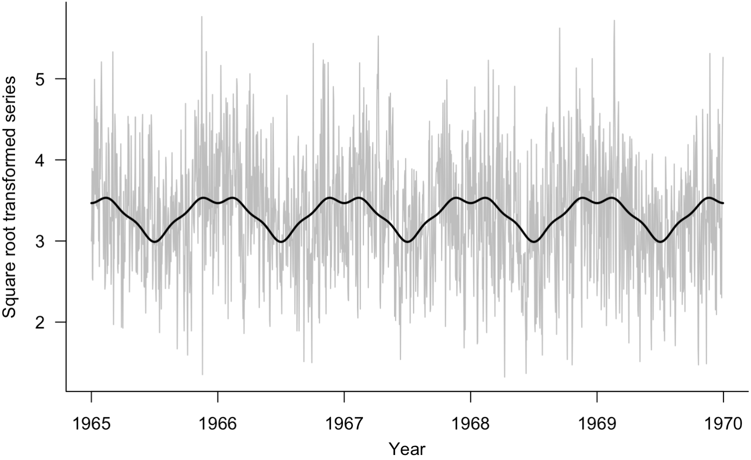
Seasonal modeling (via harmonic regression) effectively removes the oscillatory pattern in the ACF of the original series

Transform Data to Approximate Gaussian Distribution



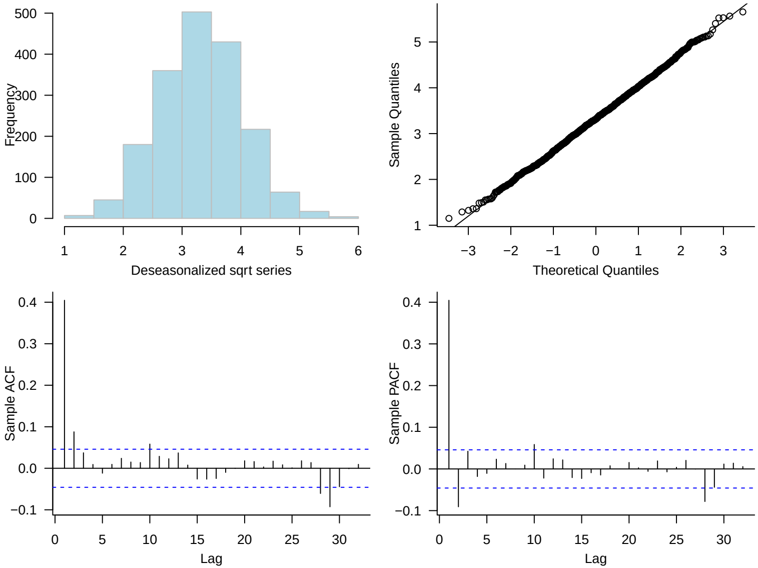
Square root transformation works! Now take the square root of the original data and deseasonalize again!

Estimating Transformed Series Seasonality



Next, we need to check if the deseasonalized series Gaussian like

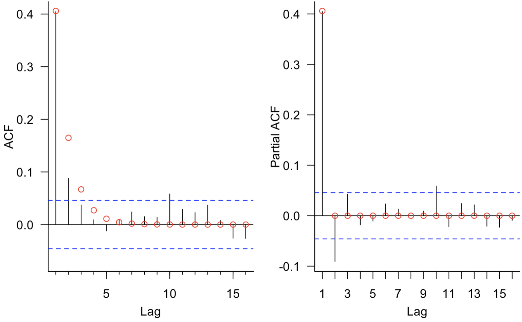
Marginal and ACF/PACF of the Deseasonalized Series



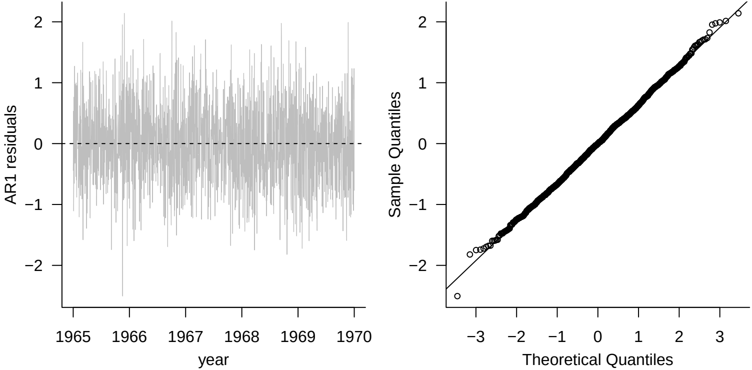
Based on ACF/PACF, which ARMA model would you choose?

Potential Model 1: AR(1)

```
> ar1.model <- arima(sqrt.rosslare.ds, order = c(1, 0, 0))
> ar1.model

Call:
arima(x = sqrt.rosslare.ds, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.4060     3.3257
s.e.  0.0214     0.0254

sigma^2 estimated as 0.4148:  log likelihood = -1787.72,  aic = 3581.43
```



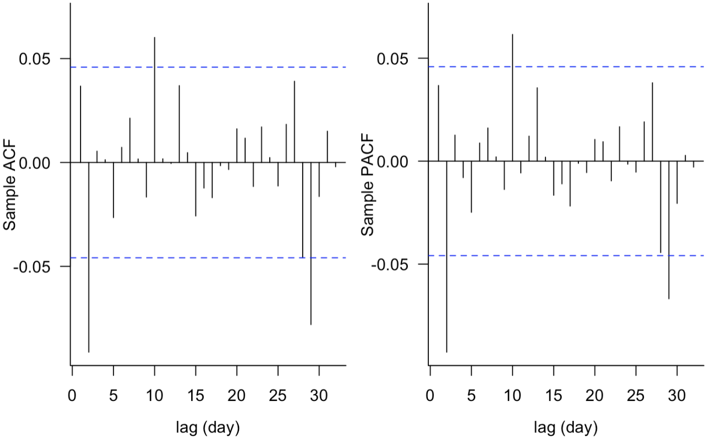
Residual Plots for the AR(1) Model



Normality assumption seems reasonable.

Next check the ACF/PACF and perform a Box test to assess if the AR(1) fit adequately account for temporal dependence strucuture

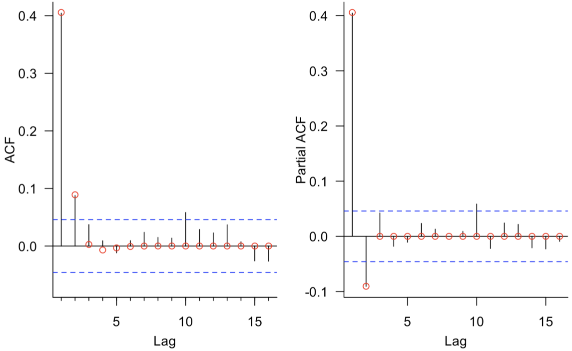
Diagnostic for the AR(1) Model



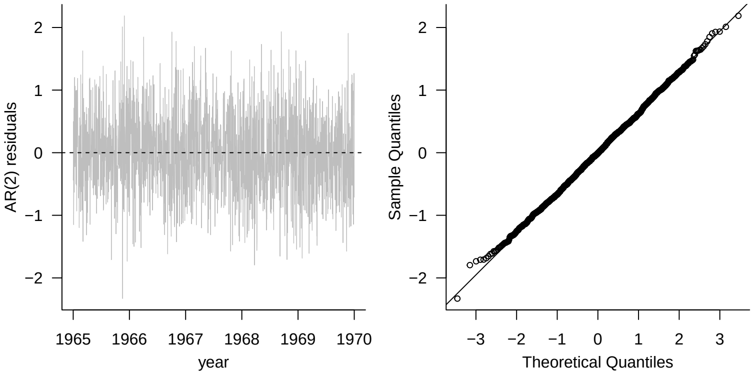

```
> Box.test(ar1.resids, lag = 32, fitdf = 1, type = "Ljung-Box")

        Box-Ljung test

data:  ar1.resids
X-squared = 53.142, df = 31, p-value = 0.00794
```

Potential Model 2: AR(2)

```
> (ar2.model <- arima(sqrt.rosslare.ds, order = c(2, 0, 0)))

Call:
arima(x = sqrt.rosslare.ds, order = c(2, 0, 0))

Coefficients:
         ar1      ar2  intercept
      0.4425  -0.0905     3.3254
s.e.  0.0233   0.0233     0.0232

sigma^2 estimated as 0.4114:  log likelihood = -1780.23,  aic = 3568.46
```



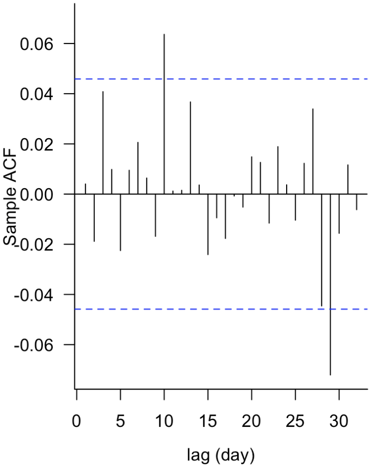
Residual Plots for the AR(2) Model

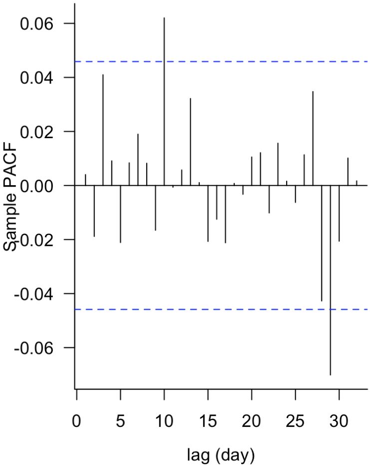


Normality assumption seems reasonable.

Next check the ACF/PACF and perform a Box test to assess if the AR(2) fit adequately account for temporal dependence strucuture

Diagnostic for the AR(2) Model

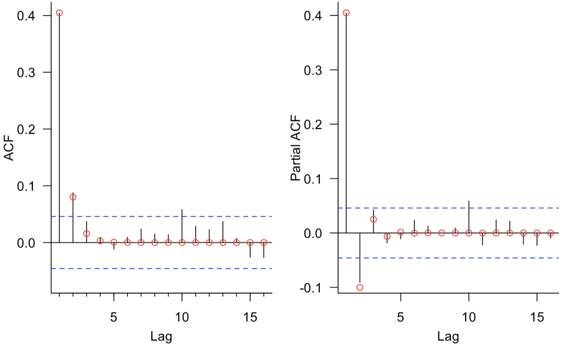




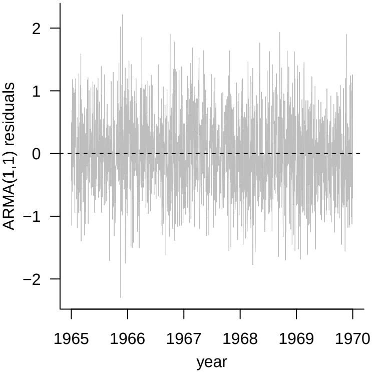

```
> Box.test(ar2.resids, lag = 32, fitdf = 2, type = "Ljung-Box")

        Box-Ljung test

data:  ar2.resids
X-squared = 36.548, df = 30, p-value = 0.1907
```

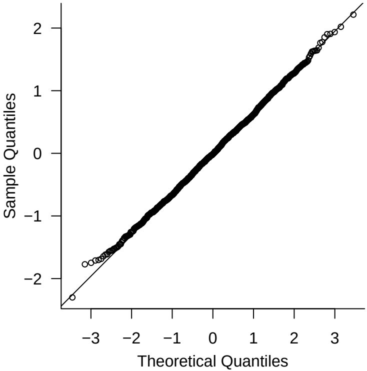
Potential Model 3: ARMA(1, 1)

```
> (arma11.model <- arima(sqrt.rosslare.ds, order = c(1, 0, 1)))

Call:
arima(x = sqrt.rosslare.ds, order = c(1, 0, 1))

Coefficients:
         ar1     ma1  intercept
      0.1978  0.2502     3.3254
s.e.  0.0556  0.0553     0.0234

sigma^2 estimated as 0.4108:  log likelihood = -1778.82,  aic = 3565.64
```



Residual Plots for the ARMA(1, 1) Model

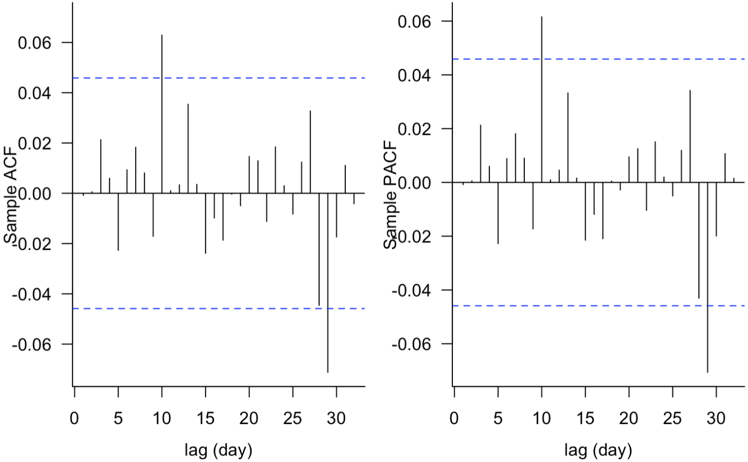




Normality assumption seems reasonable.

Next check the ACF/PACF and perform a Box test to assess if the ARMA(1, 1) fit adequately account for temporal dependence strucuture

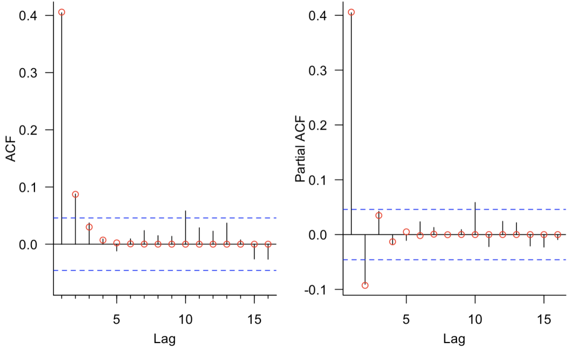
Diagnostic for the ARMA(1, 1) Model



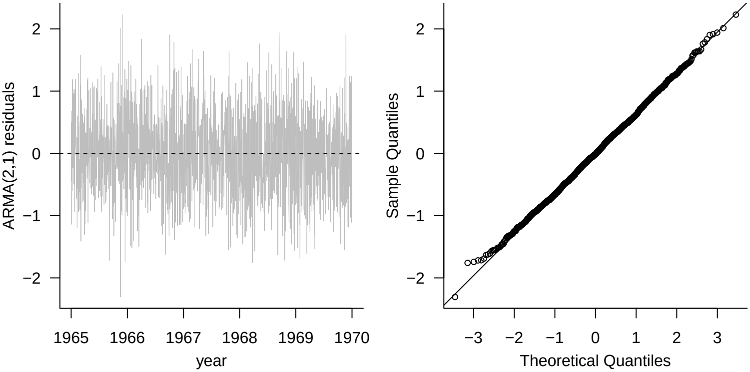

```
> Box.test(arma11.resids, lag = 32, fitdf = 2, type = "Ljung-Box")

        Box-Ljung test

data:  arma11.resids
X-squared = 32.757, df = 30, p-value = 0.3332
```

Potential Model 4: ARMA(2, 1)

```
> (arma21.model <- arima(sqrt.rosslare.ds, order = c(2, 0, 1)))

Call:
arima(x = sqrt.rosslare.ds, order = c(2, 0, 1))

Coefficients:
         ar1     ar2     ma1  intercept
      0.0703  0.0587  0.3768     3.3253
s.e.  0.1691  0.0772  0.1663     0.0237

sigma^2 estimated as 0.4107:  log likelihood = -1778.56,  aic = 3567.11
```



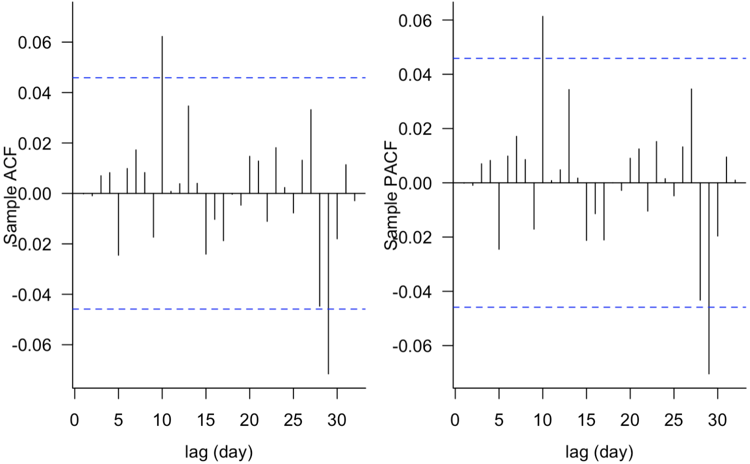
Residual Plots for the ARMA(2, 1) Model



Normality assumption seems reasonable.

Next check the ACF/PACF and perform a Box test to assess if the ARMA(2, 1) fit adequately account for temporal dependence strucuture

Diagnostic for the ARMA(2, 1) Model




```
> Box.test(arma21.resids, lag = 32, fitdf = 3, type = "Ljung-Box")

        Box-Ljung test

data:  arma21.resids
X-squared = 32.171, df = 29, p-value = 0.3124
```

Comparing Models via Information Criteria

Model	AIC	AICc
AR(1)	3583.817	3583.824
AR(2)	3570.650	3570.663
ARMA(1, 1)	3567.833	3567.847
ARMA(2, 1)	3569.319	3569.341

Which model would you pick?

Forecasting Future Wind Speeds

Question: How do we predict wind speeds on the original scale, including the seasonality that was previously estimated?

- Suppose we want to predict the next 7 days of wind speed values. We base our forecasts on the chosen ARMA(1,1) model.
- We need to reverse the order of our modeling process: ⇒ forecast under the transformed scale → add the estimated seasonal component → back-transform to the original scale.

Forecasting Future Wind Speeds, continued

- The forecasts for the next 7 days of deseasonalized square root values are:

```
> round(sqrt.rosslare.forecast$pred, 3)
Time Series:
Start = c(1970, 1)
End = c(1970, 7)
Frequency = 365
[1] 3.997 3.458 3.352 3.331 3.326 3.326 3.325
```

- The standard error for the forecasts are:

```
> round(sqrt.rosslare.forecast$se, 3)
Time Series:
Start = c(1970, 1)
End = c(1970, 7)
Frequency = 365
[1] 0.641 0.702 0.705 0.705 0.705 0.705 0.705
```

Forecasting future wind speeds, continued

- Next, we add back in the seasonality to get:

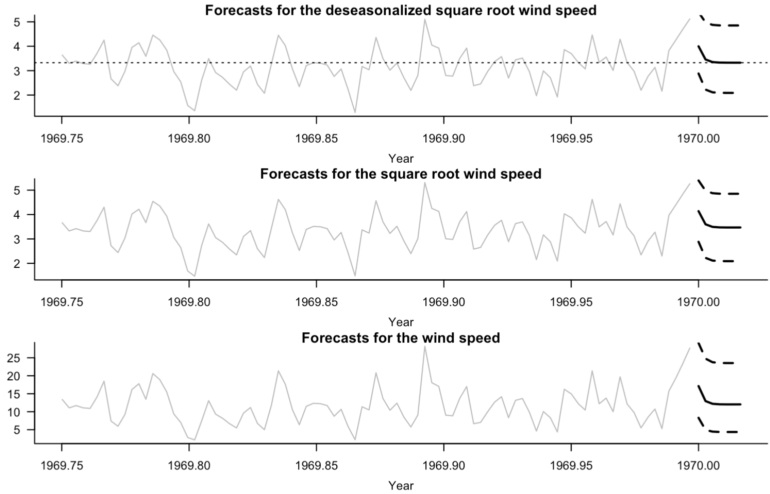
```
> adj.forecast <- fitted(harm.model)[1:h] + sqrt.rosslare.forecast$pred
> round(adj.forecast, 3)
Time Series:
Start = c(1970, 1)
End = c(1970, 7)
Frequency = 365
    1     2     3     4     5     6     7
4.139 3.600 3.494 3.473 3.470 3.470 3.470
```

- Finally, we transform back to the original scale

```
> round(adj.forecast^2, 3)
Time Series:
Start = c(1970, 1)
End = c(1970, 7)
Frequency = 365
     1      2      3      4      5      6      7
17.132 12.962 12.208 12.064 12.039 12.039 12.044
```

- To get the prediction limits, we need to transform the lower and upper prediction limits on the sqrt scale

Visualizing the Forecasts

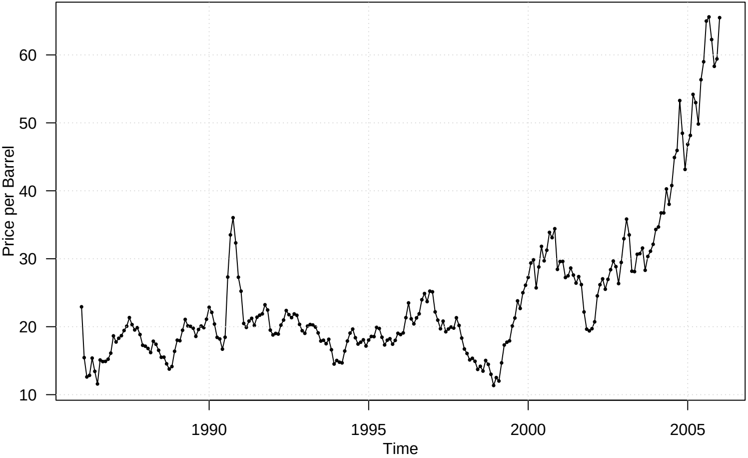


Further Questions

- What is the full model for our time series data?
- Is there a better description for the trend than just a constant term? What about alternative seasonal modeling?
- How well do we forecast? What about forecast uncertainty?

Autoregressive Integrated Moving Average (ARIMA) Models

Monthly Price of Oil: January 1986–January 2006



A stationary model does not seem to be reasonable. However, it is also not clear which (deterministic) trend model is appropriate ☹

Random Walks Revisited

Recall the random walk process

$$X_t = Z_1 + Z_2 + \cdots + Z_t = \sum_{j=1}^{t} Z_j,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

$\{X_t\}$ is a nonstationary process

- We can obtain a stationary process by differencing

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

Random Walks Revisited

Recall the random walk process

$$X_t = Z_1 + Z_2 + \cdots + Z_t = \sum_{j=1}^{t} Z_j,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

$\{X_t\}$ is a nonstationary process

- We can obtain a stationary process by differencing

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

- $\{X_t\}$ is an example of an autoregressive integrated moving average (ARIMA) process– ARIMA(0, 1, 0) process

ARIMA Models

An ARIMA model is an ARMA process after differencing

- Let d be a non-negative integer. Then X_t is an ARIMA(p, d, q) process if

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

is a causal ARMA process

ARIMA Models

An ARIMA model is an ARMA process after differencing

- Let d be a non-negative integer. Then X_t is an ARIMA(p, d, q) process if

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

is a causal ARMA process

- Let $\phi(B)$ be the AR polynomial and $\theta(B)$ be the MA polynomial. Then for $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

$$\phi(B) Y_t = \theta(B) Z_t,$$

and since $Y_t = (1 - B)^d X_t$, we have

$$\phi(B)(1 - B)^d X_t = \theta(B) Z_t$$

Example: ARIMA(1, 1, 0)

Let $\phi(z) = 1 - \phi_1 z$, $\theta(z) = 1$ and $d = 1$. For a causal stationary solution (after differencing) we need to assume $|\phi_1| < 1$. Then $\{X_t\}$ is an ARIMA (1, 1, 0) process,

$$(1 - \phi_1 B)(1 - B)X_t = Z_t,$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$

Now let $Y_t = (1 - B)X_t = X_t - X_{t-1}$, after some rearrangements we have

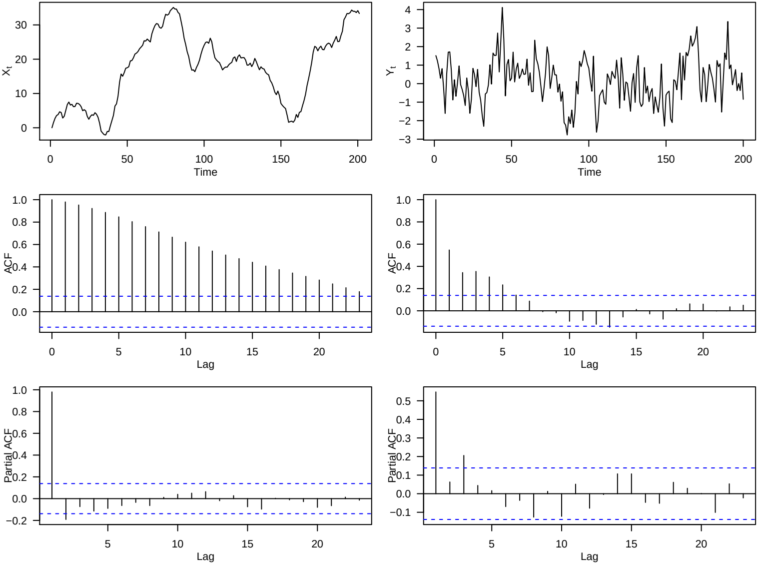
$$\begin{aligned} X_t &= X_{t-1} + Y_t \\ &= (X_{t-2} + Y_{t-1}) + Y_t \\ &\;\vdots \\ &= X_0 + \sum_{j=1}^{t} Y_j \end{aligned}$$

Thus $\{X_t\}$ is a “sort of random walk”–we cumulatively sum an AR(1) process, $\{Y_t\}$

Simulated ARIMA and Differenced ARMA Process

We simulate an ARIMA$(1,1,0)$:

$$(1-0.5B)(1-B)X_t = Z_t, \quad \{Z_t\} \sim \mathrm{N}(0,1)$$



Adding a Polynomial Trend

For $d \geq 1$, let $\{X_t\}$ be an ARIMA(p, d, q) process. Then $\{X_t\}$ satisfies the equation

$$\phi(B)(1 - B)^d X_t = \theta(B) Z_t$$

- Let μ_t be a polynomial of degree $(d - 1)$, i.e., $\mu_t = \sum_{j=0}^{d-1} a_j t^j$ for constants $\{a_j\}$

Adding a Polynomial Trend

For $d \geq 1$, let $\{X_t\}$ be an ARIMA(p, d, q) process. Then $\{X_t\}$ satisfies the equation

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t$$

- Let μ_t be a polynomial of degree $(d-1)$, i.e., $\mu_t = \sum_{j=0}^{d-1} a_j t^j$ for constants $\{a_j\}$
- Now let $V_t = \mu_t + X_t$, then

$$\begin{aligned}
\phi(B)(1-B)^d V_t &= \phi(B)(1-B)^d(\mu_t + X_t) \\
&= \phi(B)(1-B)^d \mu_t + \phi(B)(1-B)^d X_t \\
&= 0 + \phi(B)(1-B)^d X_t \\
&= \theta(B) Z_t
\end{aligned}$$

Adding a Polynomial Trend

For $d \geq 1$, let $\{X_t\}$ be an ARIMA(p, d, q) process. Then $\{X_t\}$ satisfies the equation

$$\phi(B)(1 - B)^d X_t = \theta(B) Z_t$$

- Let μ_t be a polynomial of degree $(d-1)$, i.e., $\mu_t = \sum_{j=0}^{d-1} a_j t^j$ for constants $\{a_j\}$
- Now let $V_t = \mu_t + X_t$, then

$$\begin{aligned}
\phi(B)(1-B)^d V_t &= \phi(B)(1-B)^d(\mu_t + X_t) \\
&= \phi(B)(1-B)^d \mu_t + \phi(B)(1-B)^d X_t \\
&= 0 + \phi(B)(1-B)^d X_t \\
&= \theta(B) Z_t
\end{aligned}$$

- Takeaway: ARIMA(p, d, q) are useful for modeling data with polynomial trends, due to the inherent differencing that can be used to remove trends

Steps for Modeling ARIMA Processes: Exploratory Analysis

- Plot the data, ACF, PACF and Q-Q plots
  - Check for unusual features of the data
  - Check for stationarity
  - Do we need to transform the data?
- Eliminate trend
  - Estimating the trend and removing it from the series
  - Or, differencing the series (i.e., select d in the ARIMA model)
- Plot the sample ACF/PACF for the stationary component
  - Identify candidate values of p and q

Steps for Modeling ARIMA Processes: Estimation and Model Checking

- Estimate the ARMA process parameters for the candidate models
- Check the goodness of fit: Are the time series residuals, $\{r_t\}$ a sample of *i.i.d.* noise?
- Model selection:
  - Using information criteria such as AIC and AICC
  - Test model parameters to compare between the "full" model and the "subset" model