2022 Fall

IE 313 Time Series Analysis

8. Model Diagnostics

Yongjae Lee

Department of Industrial Engineering

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Chapter 8. Model Diagnostics

- 8.1 Residual Analysis
- 8.2 Overfitting and Parameter Redundancy

Chapter 8.1

Residual Analysis

Residual analysis

residual = actual - predicted

- If the model is **correctly** specified and the parameter estimates are reasonably close to the true values, then the **residuals should have nearly the properties of white noise**
  - They should behave roughly like independent, identically distributed normal variables with zero means and common standard deviations
  - Deviations from these properties can help us discover a more appropriate model

Residual analysis

- Example
  - AR(2): $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \theta_0 + e_t$
    - Having estimated ϕ_1, ϕ_2, and θ_0, the residuals would be

$$\hat{e}_t = Y_t - \hat{\phi}_1 Y_{t-1} - \hat{\phi}_2 Y_{t-2} - \hat{\theta}_0$$

  - General ARMA (containing MA terms)
    - These models include past white noise terms, and thus, they cannot be imply represented as above
    - But we can use the inverted, infinite AR form of the model

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \pi_3 Y_{t-3} + \cdots + e_t$$

    - Then

$$\hat{e}_t = Y_t - \hat{\pi}_1 Y_{t-1} - \hat{\pi}_2 Y_{t-2} - \hat{\pi}_3 Y_{t-3} - \cdots$$

Residual analysis

- Plot of residuals
  - First diagnostic check is to inspect a plot of the residuals over time
    - If the model is adequate, we expect the plot to suggest a rectangular scatter around a zero horizontal level with no trends whatsoever
- Normality of the residuals
  - As we have seen in Chapter 3, quantile-quantile plots are an effective tool for assessing normality
- Autocorrelation of the residuals
  - To check on the independence of the noise terms in the model, we consider the sample ACF of the residuals

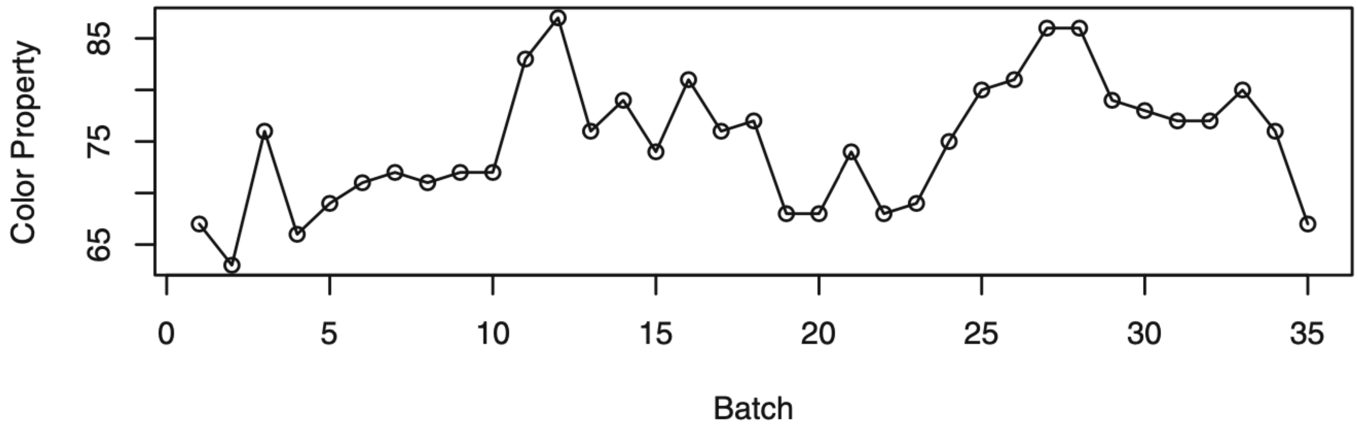
Residual analysis

- Ljung-Box test
  - In addition to looking at residual correlations at individual lags, it is useful to have a test that takes account their magnitudes as a group
  - Box and Pierce (1970)
    - $Q = n(\hat{r}_1^2 + \hat{r}_2^2 + \cdots + \hat{r}_K^2) \sim \chi^2(K - p - q)$
      - $\hat{r}_i$: sample ACF of residuals at lag k
      - n: number of samples
    - The idea is that an erroneous model would tend to inflate Q
  - Ljung and Box (1978)
    - Proposed a modified statistic that is more appropriate for typical sample sizes
    - $Q_* = n(n+2)\left(\frac{\hat{r}_1^2}{n-1} + \frac{\hat{r}_2^2}{n-2} + \cdots + \frac{\hat{r}_K^2}{n-K}\right) \sim \chi^2(K - p - q)$

Chemical process color property series

- Original series

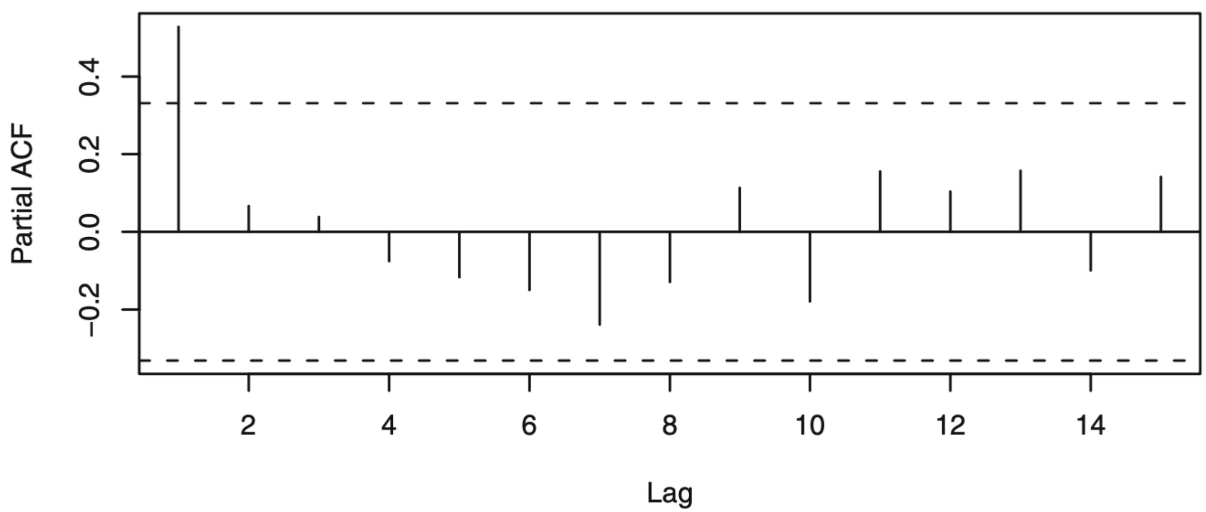
Exhibit 1.3 **Time Series Plot of Color Property from a Chemical Process**



Chemical process color property series

- Sample ACF

Exhibit 6.26 Sample Partial ACF for the Color Property Series

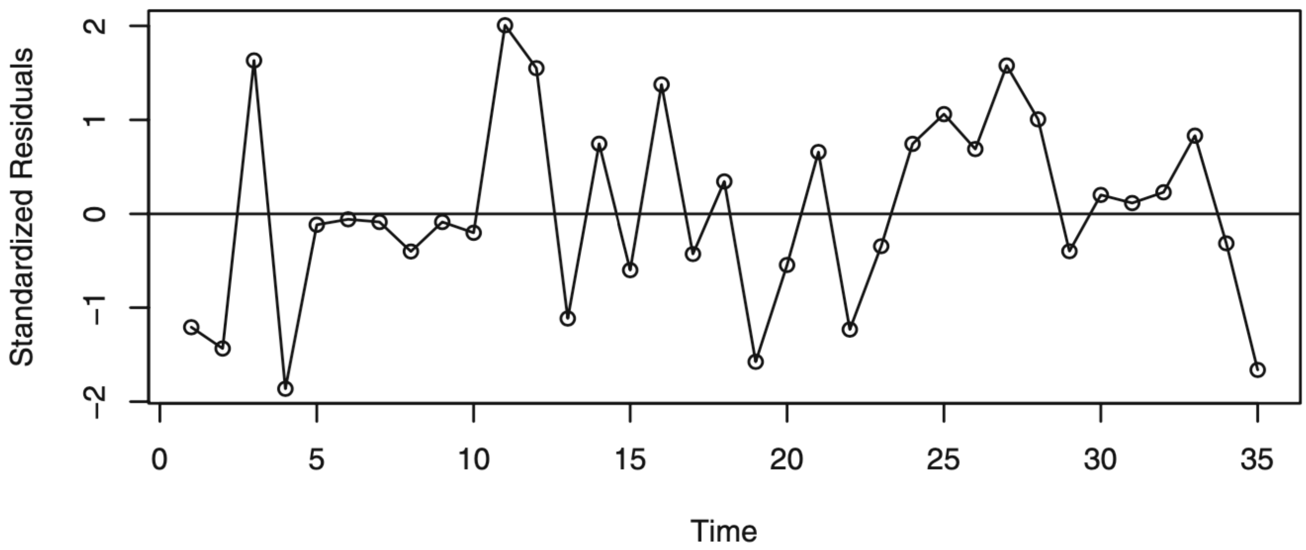


  – The sample PACF clearly suggest that an **AR(1)** model is worthy of first consideration
  – As always, our specified models are tentative and subject to modification during the model diagnostics stage of model building

Chemical process color property series

- Plot of the residuals

Exhibit 8.1 Standardized Residuals from AR(1) Model of Color

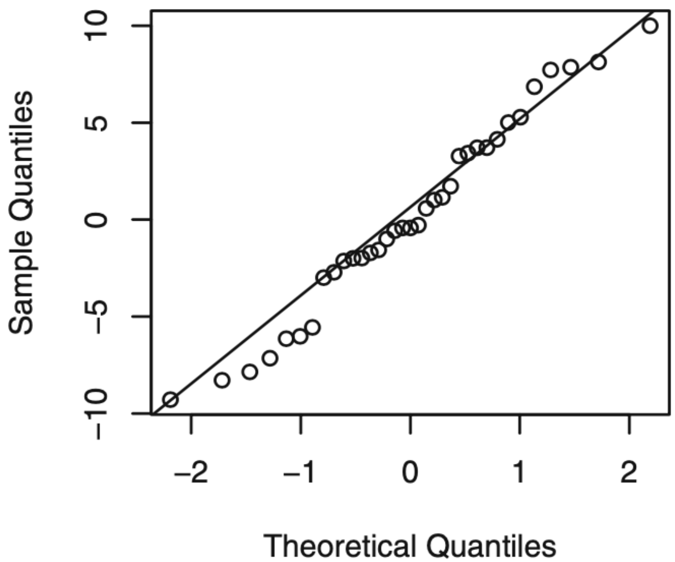


– Seem to have no trends

Chemical process color property series

- Normality of the residuals

Exhibit 8.4 **Quantile-Quantile Plot: Residuals from AR(1) Color Model**

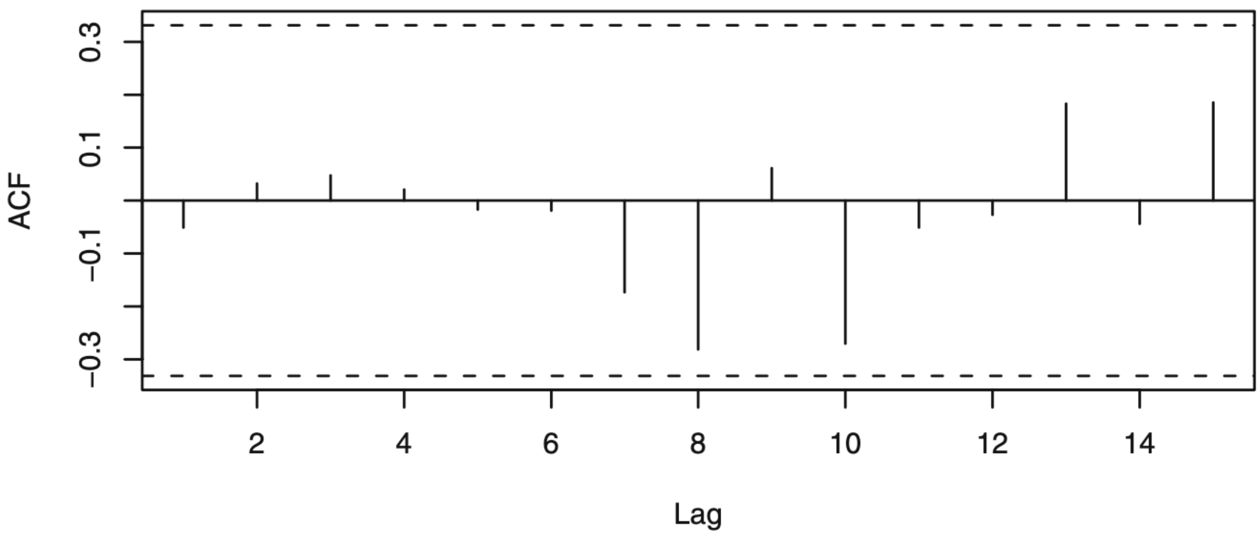


– Seem to follow the straight line fairly closely

Chemical process color property series

- Autocorrelation of the residuals

Exhibit 8.9 Sample ACF of Residuals from AR(1) Model for Color

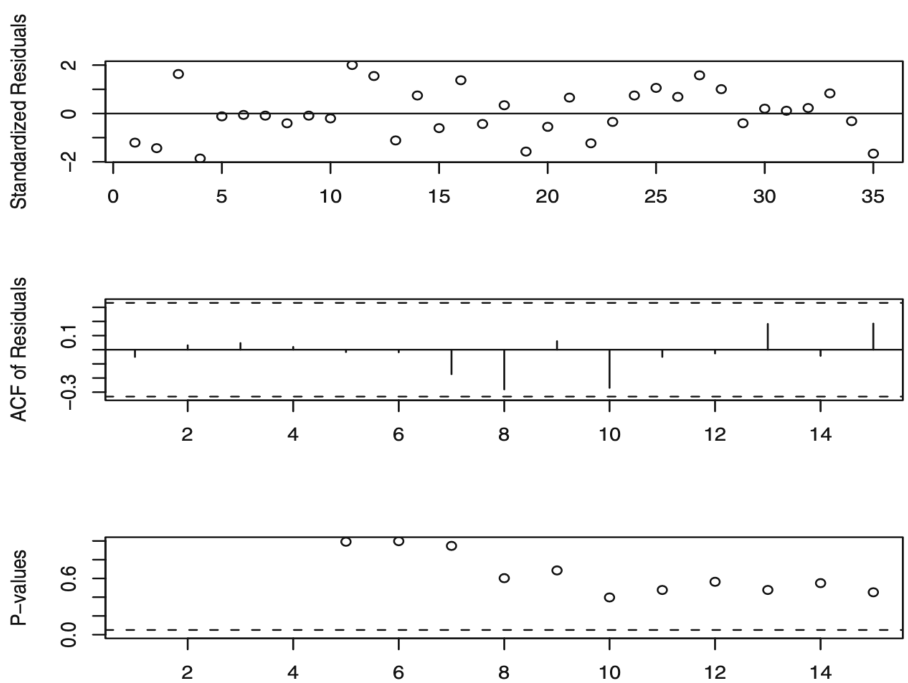


– No evidence of autocorrelation in the residuals

Chemical process color property series

- Ljung-Box test

Exhibit 8.12 Diagnostic Display for the AR(1) Model of Color Property

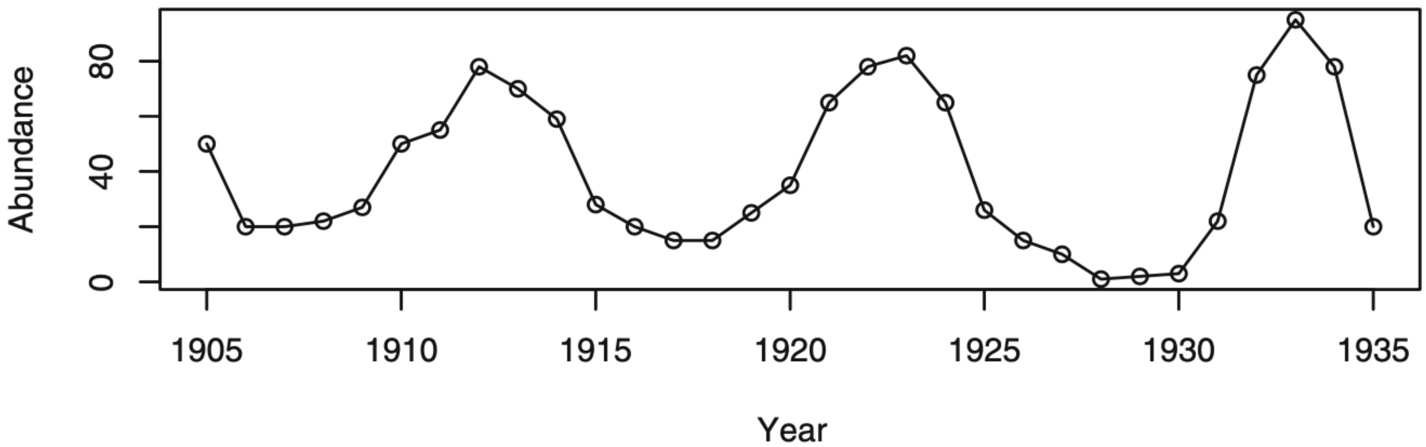


– The last figure shows the p–values for the Ljung-Box test statistic for different values of K
  - We have no problem for ACF of residuals

Abundance of Canadian hare

- Original series

Exhibit 1.5 Abundance of Canadian Hare



Abundance of Canadian hare

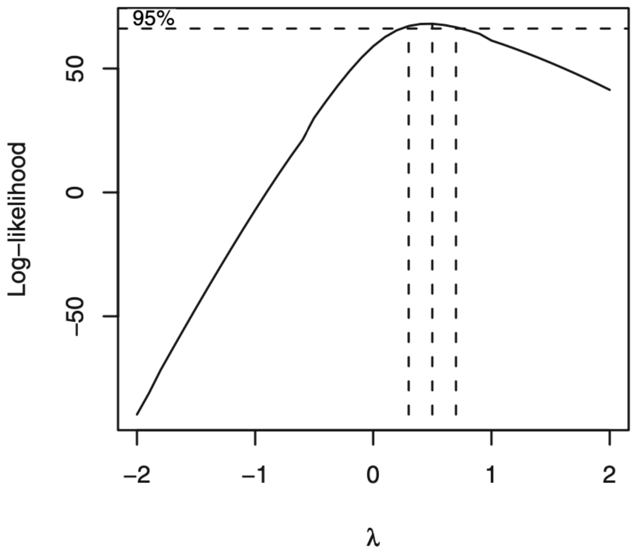
- Power transformation

Power transformations

- A more general **power transformation** was introduced by Box and Cox (1964)
  - For a given value of the parameter λ,

$$g(x) = \begin{cases} \dfrac{x^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \log x & \text{for } \lambda = 0 \end{cases}$$

Exhibit 6.27 Box-Cox Power Transformation Results for Hare Abundance

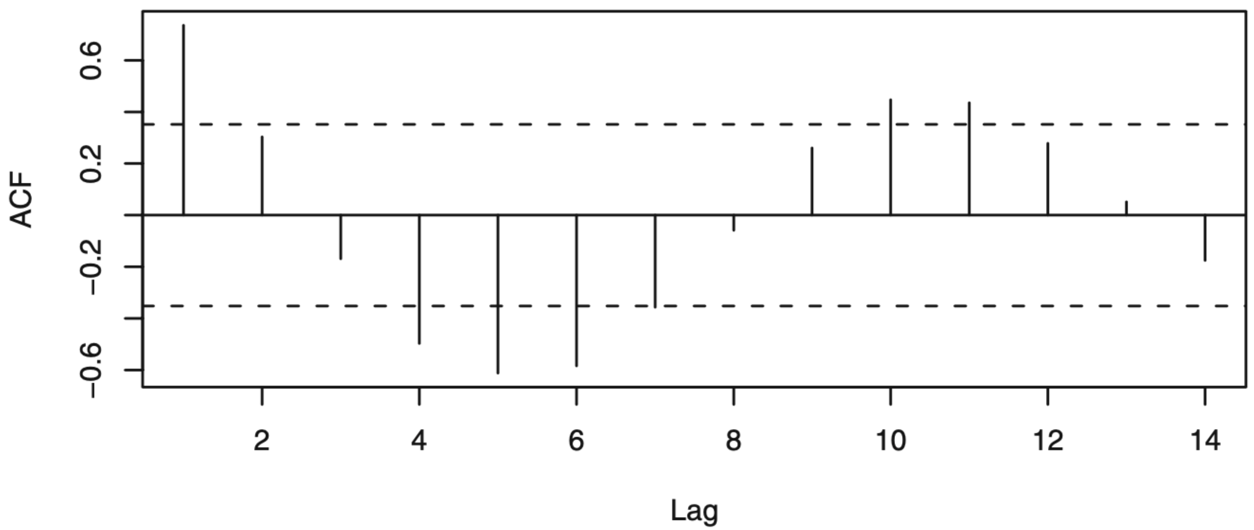


- It has been suggested in the literature that a transformation might be used to produce a good model for these data
  - Exhibit 6.27 displays the log-likelihood as a function of the power parameter λ
  - The maximum occurs at $\lambda = 0.4$, but a **square root** transformation ($\lambda = 0.5$) is well within the confidence interval for λ

Abundance of Canadian hare

- Sample ACF

Exhibit 6.28 Sample ACF for Square Root of Hare Abundance

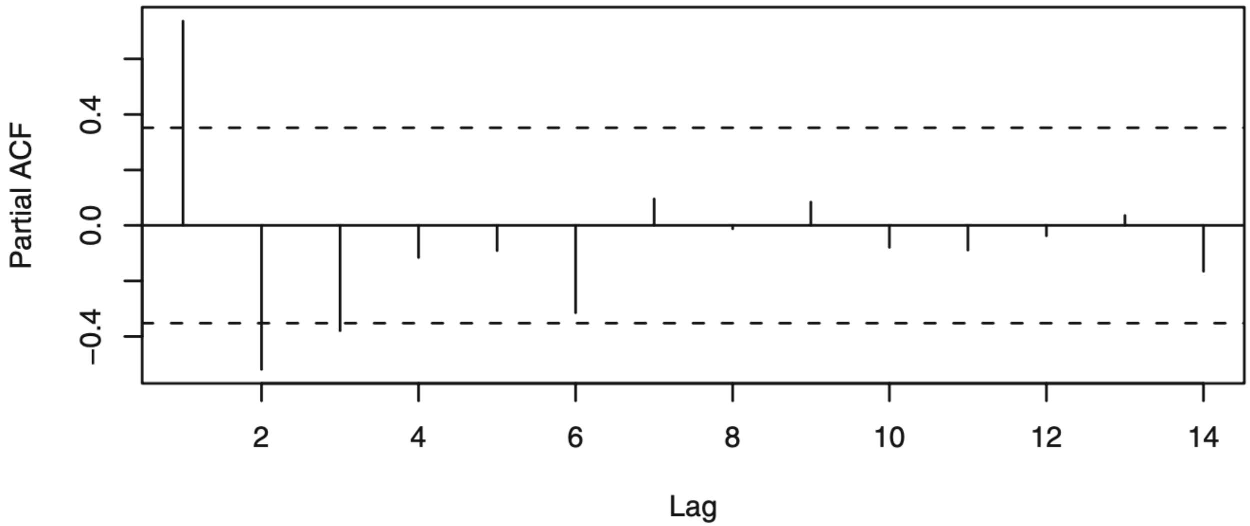


– The sample ACF of the *square root transformed* series has a fairly strong lag 1 ACF, but there is a strong indication of **dampened oscilatory behavior**

Abundance of Canadian hare

- Sample PACF

Exhibit 6.29 Sample Partial ACF for Square Root of Hare Abundance

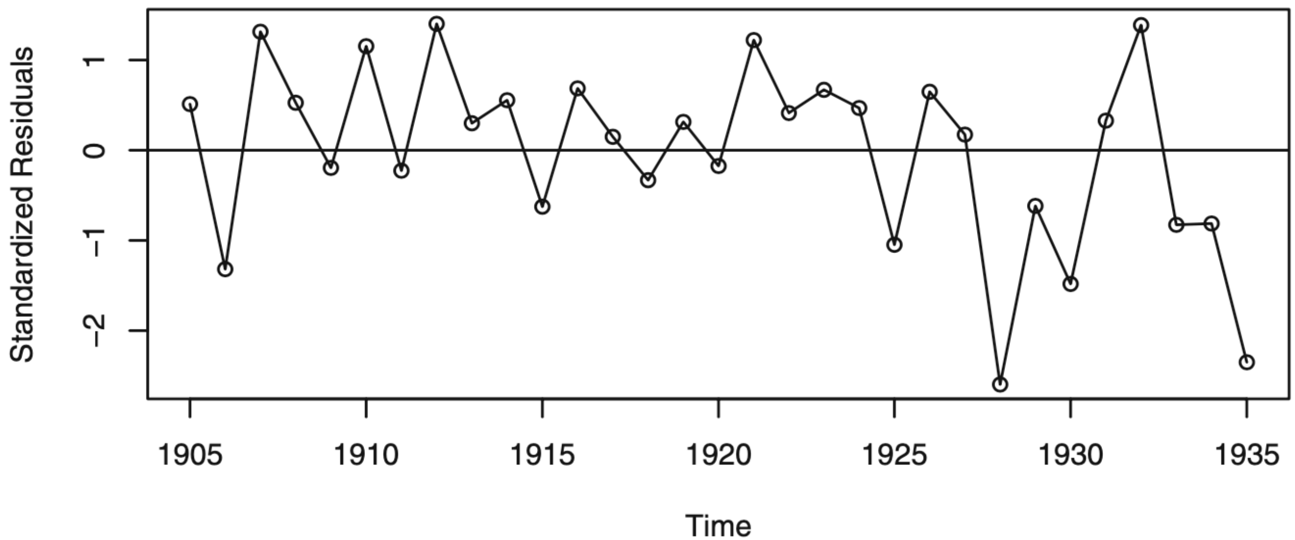


– The sample PACF of square root transformed series gives strong evidence to support **AR(2)** or possibly an **AR(3)**

Abundance of Canadian hare

- Plot of residuals

Exhibit 8.2 Standardized Residuals from AR(3) Model for Sqrt(Hare)

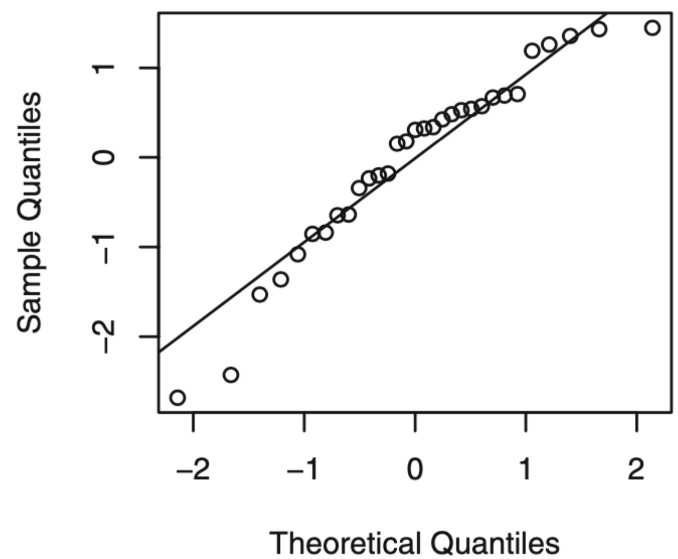


– Reduced variation in the middle and increased variation near the end

Abundance of Canadian hare

- Normality of the residuals

Exhibit 8.5 Quantile-Quantile Plot: Residuals from AR(3) for Hare

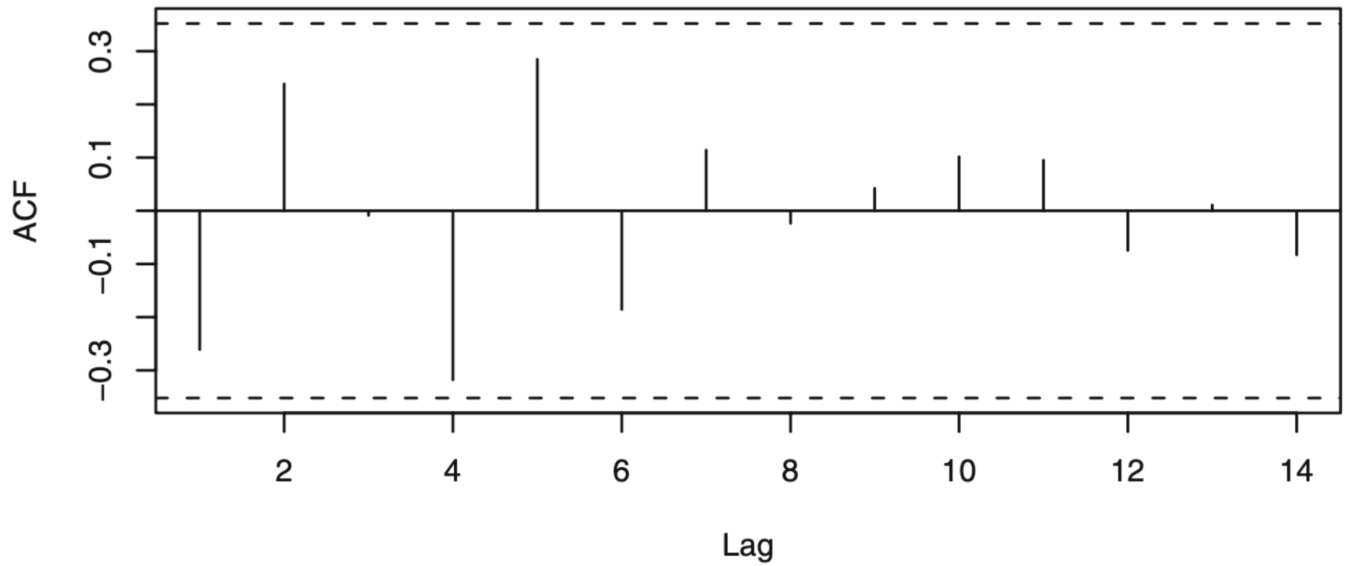


– Extreme values look suspect. But the sample is small (n=31)

Abundance of Canadian hare

- Autocorrelation of the residuals

Exhibit 8.10 Sample ACF of Residuals from AR(2) Model for Hare

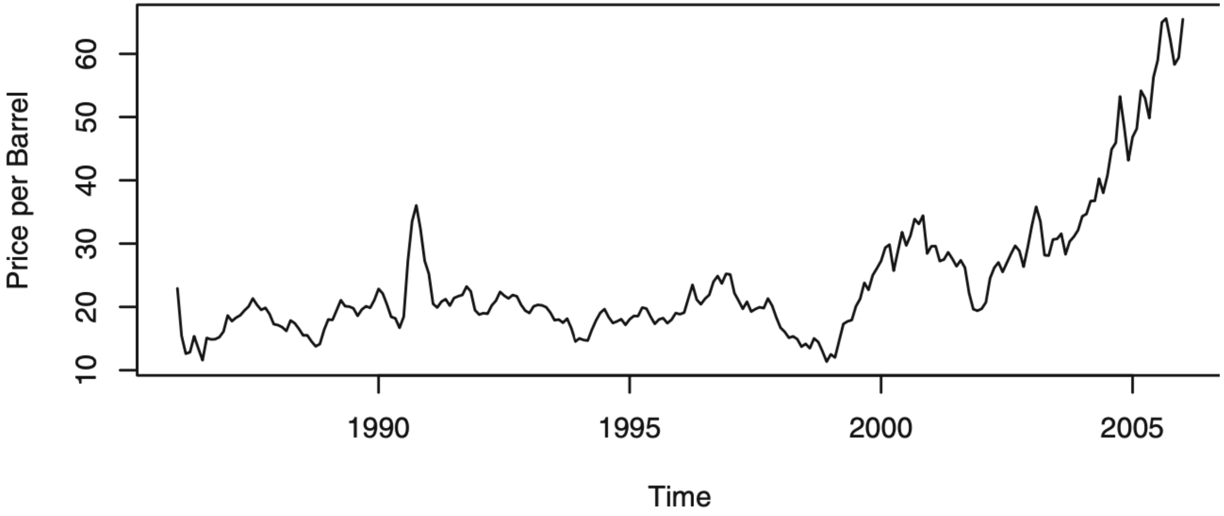


– It does not show statistically significant evidence of nonzero autocorrelation in residuals

Oil price series

- Original series

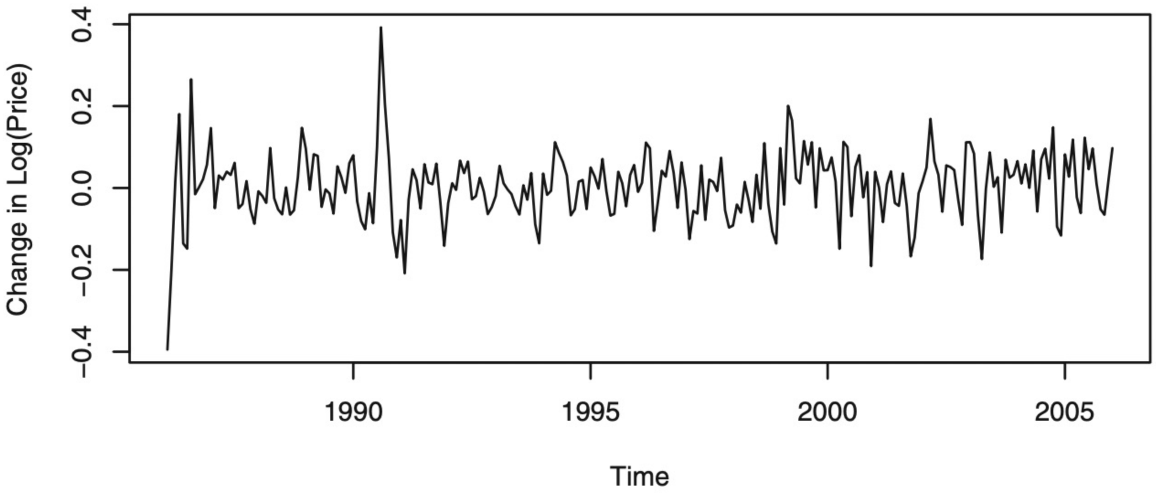
Exhibit 5.1 **Monthly Price of Oil: January 1986–January 2006**



Oil price series

- First difference of logarithms

Exhibit 5.4 The Difference Series of the Logs of the Oil Price Time



Oil price series

- EACF table

Exhibit 6.30 Extended ACF for Difference of Logarithms of Oil Price Series

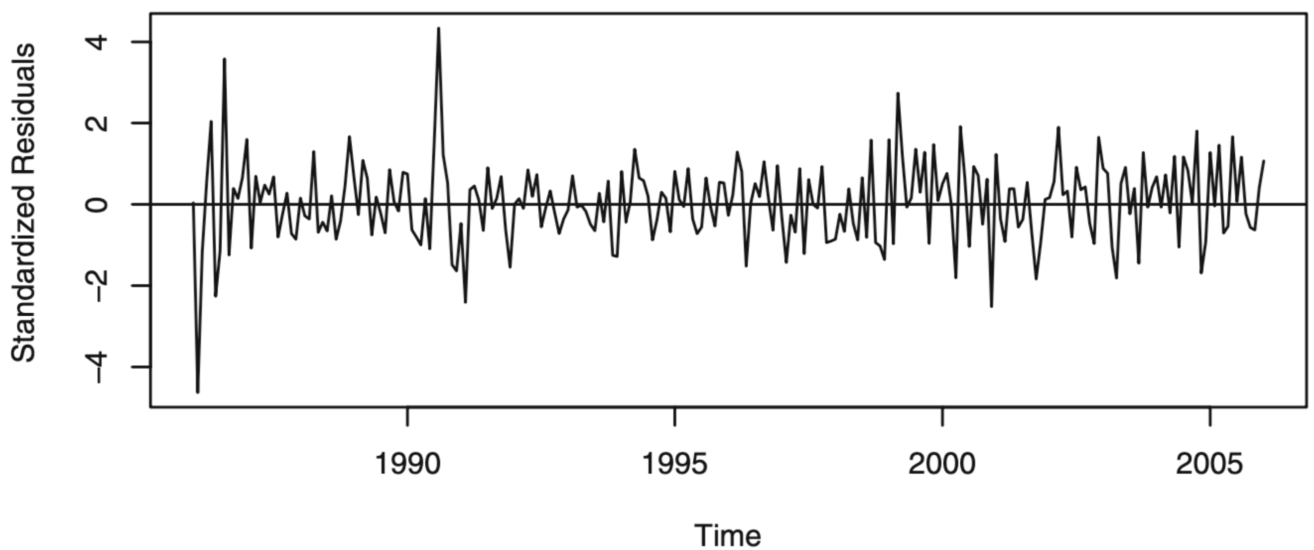
AR / MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	o	o	o	o	o	o	o	o	o	o	o	o	o
1	x	x	o	o	o	o	o	o	o	o	x	o	o	o
2	o	x	o	o	o	o	o	o	o	o	o	o	o	o
3	o	x	o	o	o	o	o	o	o	o	o	o	o	o
4	o	x	x	o	o	o	o	o	o	o	o	o	o	o
5	o	x	o	x	o	o	o	o	o	o	o	o	o	o
6	o	x	o	x	o	o	o	o	o	o	o	o	o	o
7	x	x	o	x	o	o	o	o	o	o	o	o	o	o

  – EACF table suggests that an **ARMA model with p = 0 and q = 1 (i.e., MA(1))** would be appropriate for the first difference of logs of oil price series

Oil price series

- Plot of the residuals

Exhibit 8.3 Standardized Residuals from Log Oil Price IMA(1,1) Model

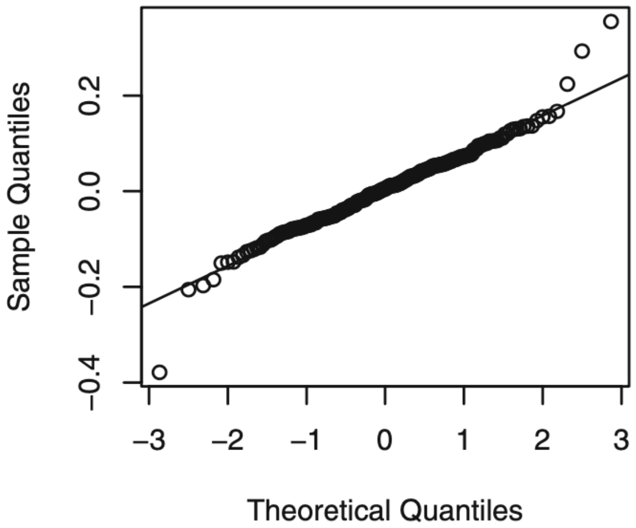


  – At least 2 or 3 residuals early in the series with magnitudes larger than 3 (very unusual in a standard normal distribution)

Oil price series

- Normality of the residuals

Exhibit 8.6 Quantile-Quantile Plot: Residuals from IMA(1,1) Model for Oil

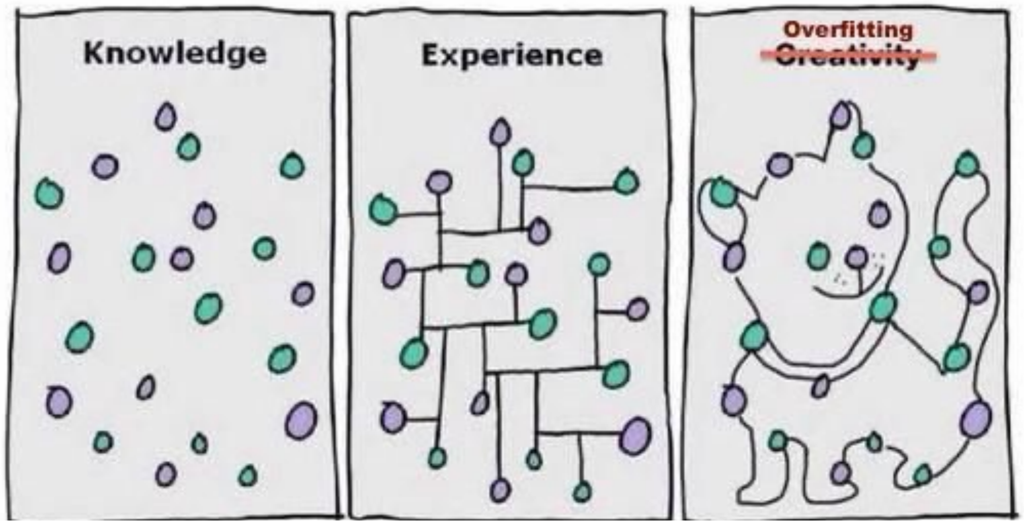


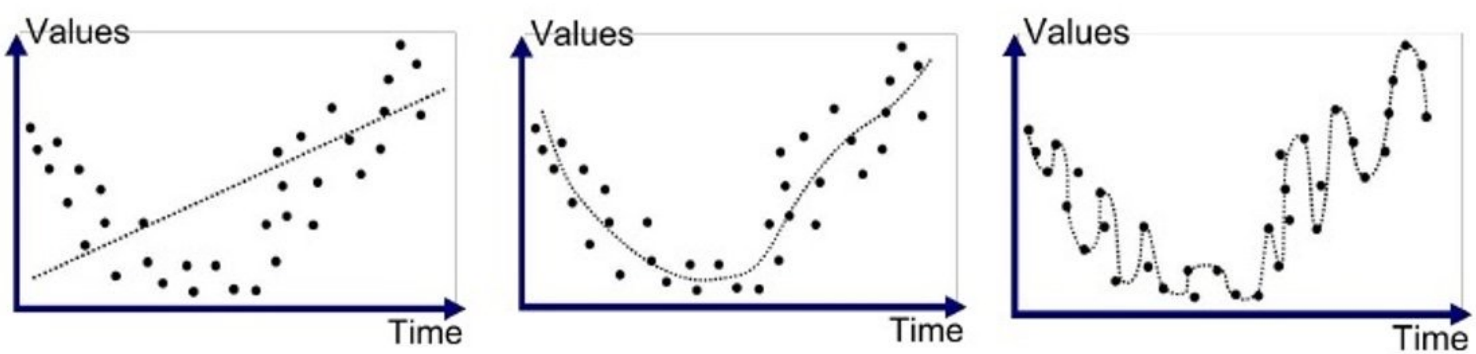
  – Outliers are quite prominent

Chapter 8.2

Overfitting and Parameter Redundancy

Overfitting





Underfitted | Good Fit/Robust | Overfitted

Overfitting

- Our second basic diagnostic tool is that of overfitting
  - After specifying and fitting what we believe to be an adequate model, we fit a slightly more general model
  - That is, a model “close by” that contains the original model as a spacial case
    - For example, if an AR(1) model seems appropriate, we might overfit with an AR(2) model
    - The original AR(1) model would be confirmed if
      - The estimate of the additional parameter ϕ_2 is not significantly different from zero
      - The estimate for the parameter in common ϕ_1 does not change significantly from their original estimates

Chemical process color property series

- **To test AR(1) we may overfit data with AR(2) or ARMA(1,1)**

Exhibit 8.13 AR(1) Model Results for the Color Property Series

Coefficients:†	ar1	Intercept‡
	0.5705	74.3293
s.e.	0.1435	1.9151

sigma^2 estimated as 24.83: log-likelihood = -106.07, AIC = 216.15

- Seem close enough
- Seem close enough

Exhibit 8.14 AR(2) Model Results for the Color Property Series

Coefficients:	ar1	ar2	Intercept
	0.5173	0.1005	74.1551
s.e.	0.1717	0.1815	2.1463

sigma^2 estimated as 24.6: log-likelihood = -105.92, AIC = 217.84

- Not significantly different from zero

Exhibit 8.15 Overfit of an ARMA(1,1) Model for the Color Series

Coefficients:	ar1	ma1	Intercept
	0.6721	−0.1467	74.1730
s.e.	0.2147	0.2742	2.1357

sigma^2 estimated as 24.63: log-likelihood = -105.94, AIC = 219.88

- Not significantly different from zero

Parameter redundancy

- Any ARMA(p,q) model can be considered as a special case of a more general ARMA model with the additional parameters equal to zero

- However, when generalizing ARMA models, we must be aware of the problem of **parameter redundancy** or **lack of identifiability**

Parameter redundancy

- Consider an ARMA(1,2) model
    - $Y_t = \phi Y_{t-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
  - Now replace t by $t - 1$
    - $Y_{t-1} = \phi Y_{t-2} + e_{t-1} - \theta_1 e_{t-2} - \theta_2 e_{t-3}$
  - If we multiply both sides of the second equation by any constant c and then subtract it from the first one, we have
    - $Y_t - (\phi + c)Y_{t-1} + \phi c Y_{t-2} = e_t - (\theta_1 + c)e_{t-1} - (\theta_2 - \theta_1 c)e_{t-2} + \theta_2 c e_{t-3}$
  - This apparently defines an ARMA(2,3) process. But notice that the AR and MA characteristic polynomials have factorizations
    - $1 - (\phi + c)x + \phi c x^2 = (1 - \phi x)(1 - cx)$
    - $1 - (\theta_1 + c)x - (\theta_2 - c\theta_1)x^2 + c\theta_2 x^3$
      $= (1 - \theta_1 x - \theta_2 x^2)(1 - cx)$
  - Y_t satisfies the ARMA(2,3) model with arbitrary parameteres (that depend on any random choice of c)

Chemical process color property series

Exhibit 8.16 Overfitted ARMA(2,1) Model for the Color Property Series

Coefficients:	**ar1**	**ar2**	**ma1**	**Intercept**
	0.2189	0.2735	0.3036	74.1653
s.e.	2.0056	1.1376	2.0650	2.1121

sigma^2 estimated as 24.58: log-likelihood = −105.91, AIC = 219.82

- Note that we have seen that an AR(1) model fits quite well for this data
- Suppose if we try an ARMA(2,1) model
  - Even though the estimate of σ_e^2 and the log-likelihood and AIC values are not too far from their best balues, the estimates of ϕ_1, ϕ_2, and θ are way off, and none would be considered different from zero statistically

Fitting and overfitting models

- The implications for fitting and overfitting models are as follows:
  1. Specify the original model carefully. If a simple model seems at all promising, check it out before trying a more complicated model
  2. When overfitting, do not increase the orders of both the AR and MA parts of the model simultaneously
  3. Extend the model in difrections suggested by the analysis of the residuals. For example, if after fitting an MA(1) model, substantial correlation remains at lag 2 in the residuals, try an MA(2), not an ARMA(1,1)