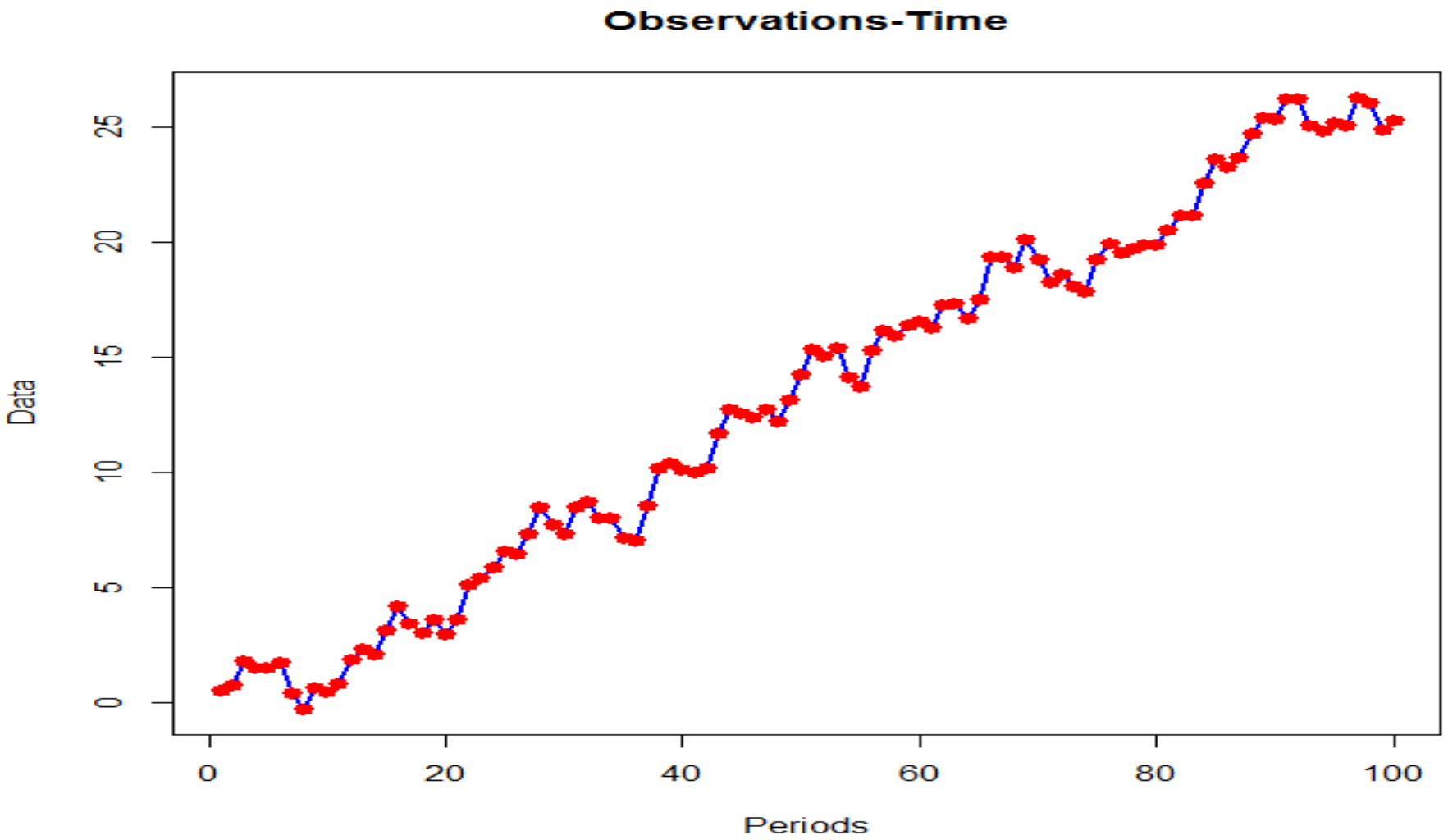
Final Project01: Data Plot



Please write down your arguments makes them clear and give necessary formulas. Cover style includes that A4 paper, and each student's name, department and id. Show clearly the question number of every answer.

Consider the following model:

$$(1-\phi_1 B-\cdots-\phi_3 B)(y_t-\mu_t)=\varepsilon_t(1-\theta_1 B-\cdots-\theta_3 B), t=1,\ldots,n,$$

where $\varepsilon_t \underset{\sim}{iid} N(0,\sigma^2)$.

Q(1). Model selection

(1) μ_t : $\{1, t, t^2, t^3, \sin t, \cos t, \ldots, \sin 4t, \cos 4t\}$.

(2) $ARMA(p,q)$: $p \times q \in \{0,1,2,3\} \times \{0,1,2,3\}$, and $p+q \geq 1$.

Let $\hat{\eta}$ be the coefficients of the selected model in Q(1).

Q(2). What are $\hat{\eta}$ and the limiting variance and covariance matrix, $\lim_{n\to\infty} Var(\hat{\eta})$? Write down formulas, computing values by some softwares, and compare them.

(Comparsion Sample)

Table 1. parameters estimates and standard derivations.

	$\hat{\beta}_1$	$\hat{\beta}_2$	...
$\hat{\boldsymbol{\beta}}$			
$sd(\hat{\boldsymbol{\beta}})$			
$\hat{\boldsymbol{\beta}}_R$			
$sd(\hat{\boldsymbol{\beta}})_R$			

Q(3). Demonstrate how to predict $y_{n+i}, i = 1, \ldots, 5$.

$\left(\sum_{i=1}^{5}(\hat{y}_{n+i} - y_{n+i})^2\right)$

Some codes

```
install.packages("MuMIn"); install.packages("utils") # some useful packages
#####################   generating arma data ##############################
n.sel = 100 ; beta0 = c( 0.1, -0.12, 0.25, -0.21 ) ; t.x = 1 : n.sel
yy = cbind( 1, t.x, sin(t.x), cos(3*t.x) )%*%beta0
     + arima.sim(list( order = c(1,0,2), ar = c(0.2), ma = c(0.1,0.3) ),sd=0.1, n = n.sel)
plot( yy, type = "l", ylab= "Data", xlab = "Periods",col="blue", main =
"Observations-Time(Sample)", lwd=2 ) ; points(yy, col="red",pch=20, lwd=5)
#------model selection to mean and p & q of arma-------------------------
dredge( , rank="" ); arima( , order = c( p ,0, q ) )
```

Seasonal Time Series Analysis

黃灝勻 f01323069@ntu.edu.tw

Seasonal Time Series Analysis

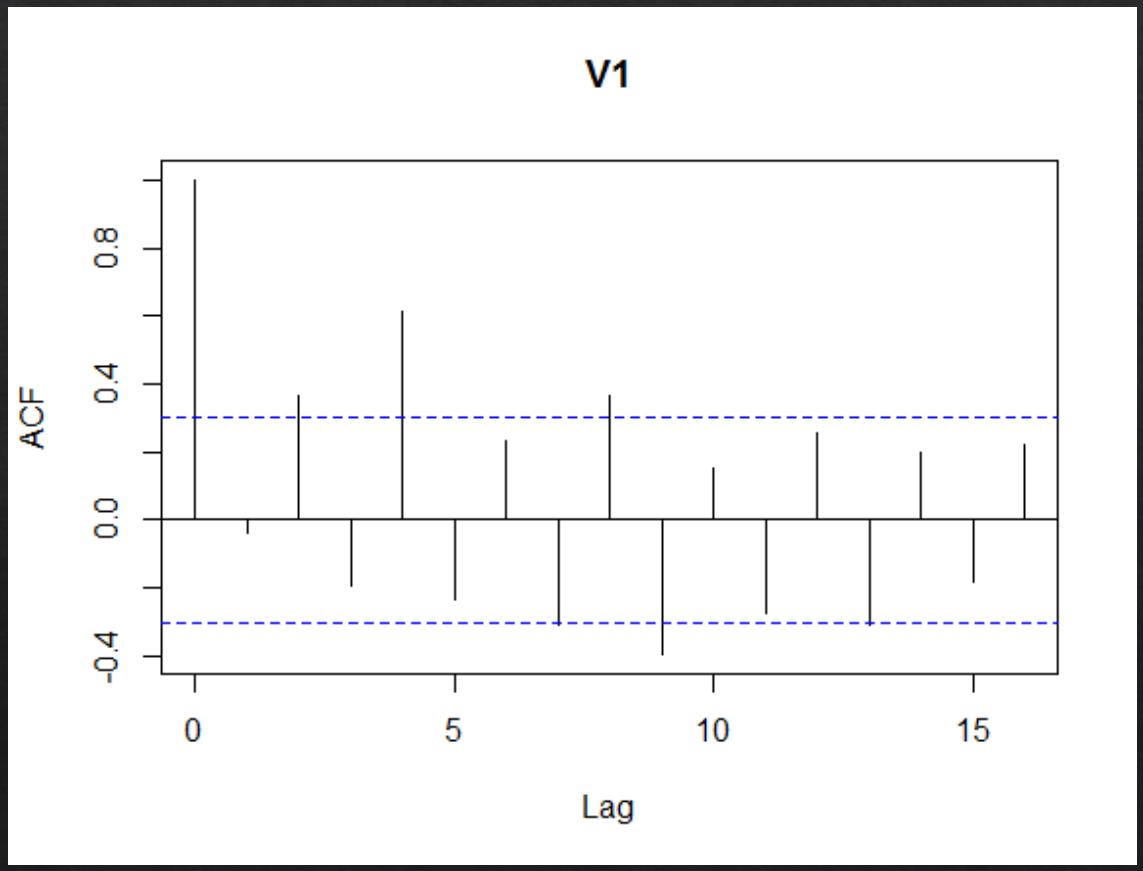
- Observations on sales of Illinois Power

- R code:

```
x=read.csv("IPQS77.csv", header = F, sep=",")
x=ts(x)
```

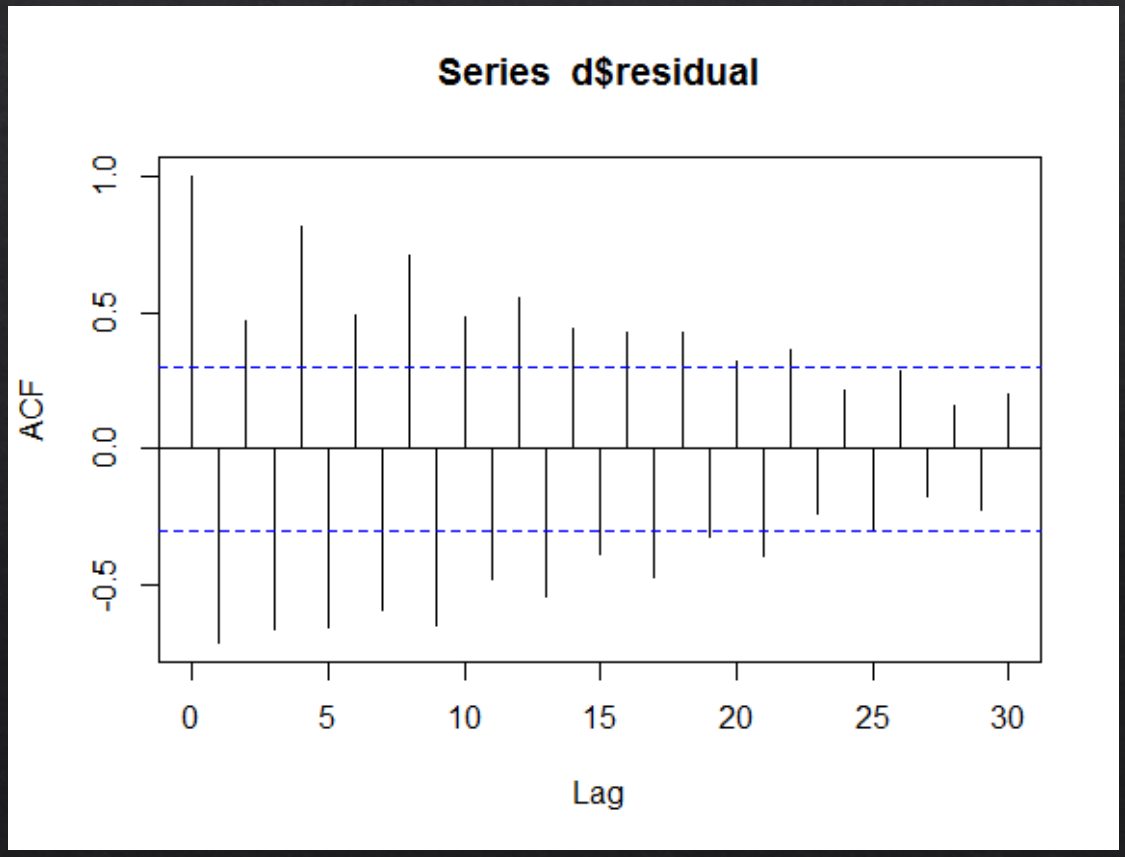
ACF of data

- acf(x)
- Dotted line $\pm 2/\sqrt{n}$ (White Noise)
- Series persistently large.



ACF after Differenced 1
$(1 - B)X_t$

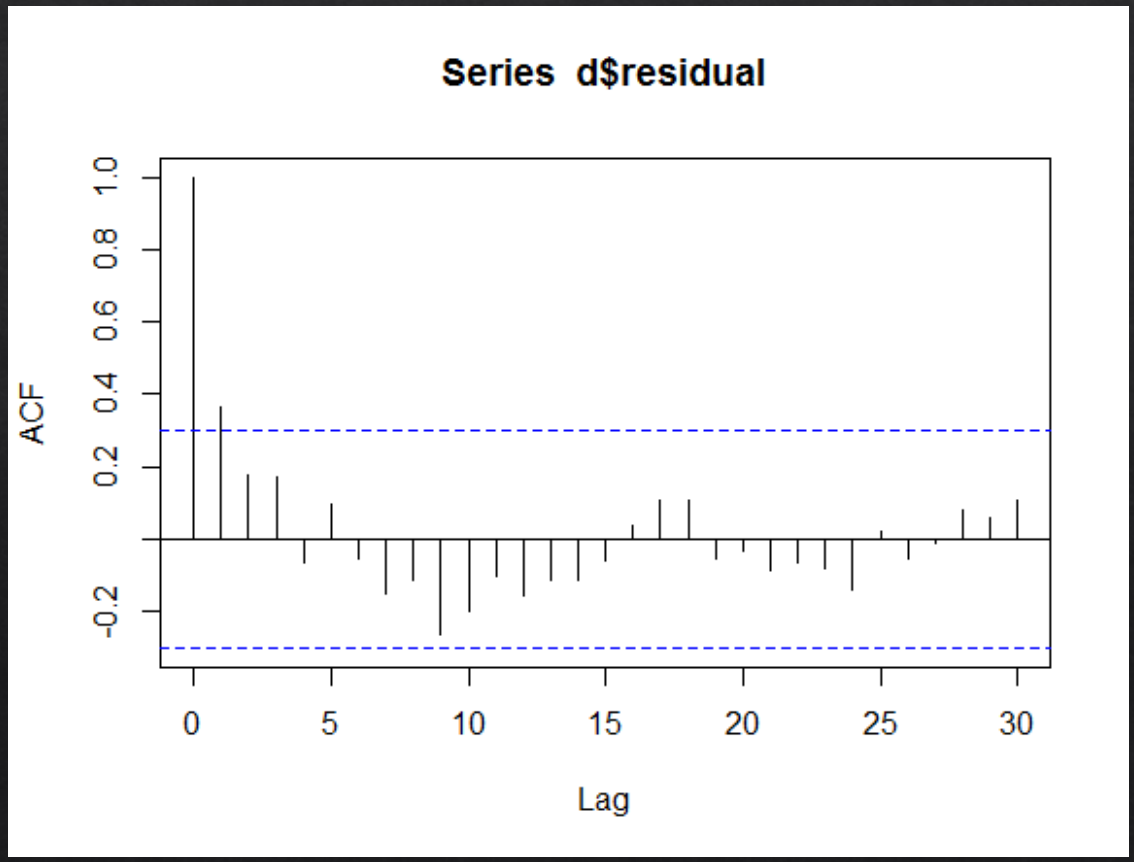
- d=arima(x,order=c(0,1,0))
- acf(d$residual,lag=30)
- Large at lag 4, 8, …



ACF after Seasonal Differenced 1 at 4

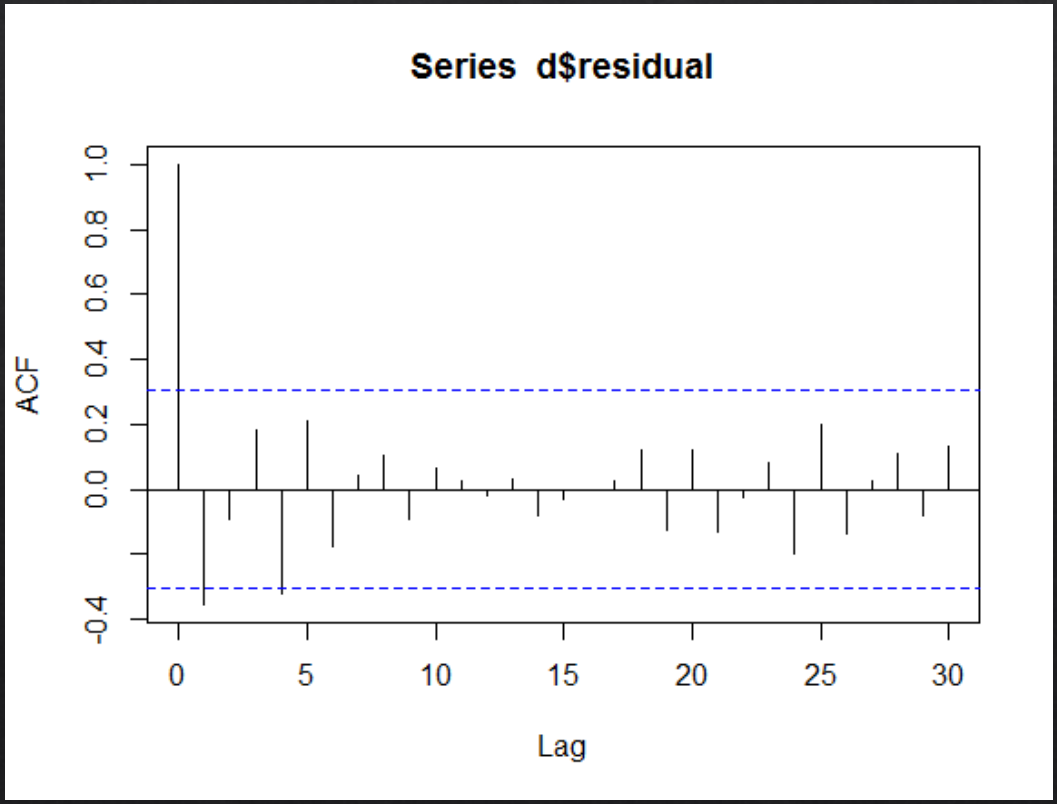
$$(1 - B^4)X_t$$

- d=arima(x,order=c(0,0,0),seasonal=list(order=c(0,1,0),period=4))
- acf(d$residual,lag=30)
- Decay smoothly.



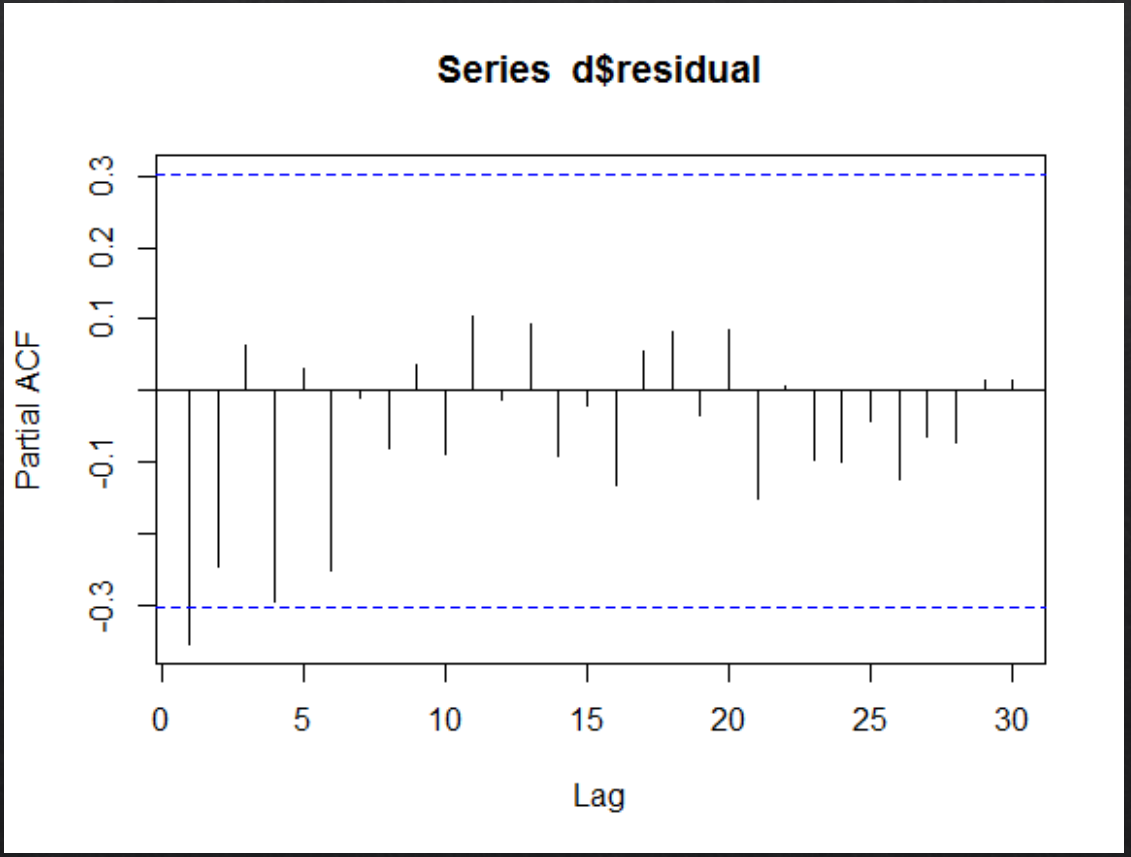
ACF of D1 and Seasonal D1 at 4
$(1-B)(1-B^4)X_t$

- d=arima(x,order=c(0,1,0),seasonal=list(order=c(0,1,0),period=12))
- acf(d$residual,lag=30)
- Large at lag 1 and lag 4



PACF of D1 and Seasonal D1 at 4

- acf(d$residual,type="partial",lag=30)
- Dotted line $\pm 2/\sqrt{n}$
  (White Noise)
- Large at lag 1, 2 and 4



Determine the Order

d1=arima(x,order=c(2,1,0),seasonal=list(order=c(1,1,0),period=4))

d2=arima(x,order=c(1,1,1),seasonal=list(order=c(1,1,0),period=4))

d3=arima(x,order=c(0,1,2),seasonal=list(order=c(1,1,0),period=4))

d4=arima(x,order=c(2,1,0),seasonal=list(order=c(0,1,1),period=4))

d5=arima(x,order=c(1,1,1),seasonal=list(order=c(0,1,1),period=4))

d6=arima(x,order=c(0,1,2),seasonal=list(order=c(0,1,1),period=4))

BIC(d1)=366.7871 (Smallest)

BIC(d2)=367.1254

BIC(d3)=366.8785

BIC(d4)=367.1581

BIC(d5)=367.4186

BIC(d6)=367.1465

Model

- $(1-\varphi_1 B-\varphi_2 B^2)(1-\alpha_1 B^4)(1-B)(1-B^4)X_t=a_t$
- $(1+0.389B+0.329B^2)(1+0.335B^4)(1-B)(1-B^4)X_t=a_t$

Run d1, we have

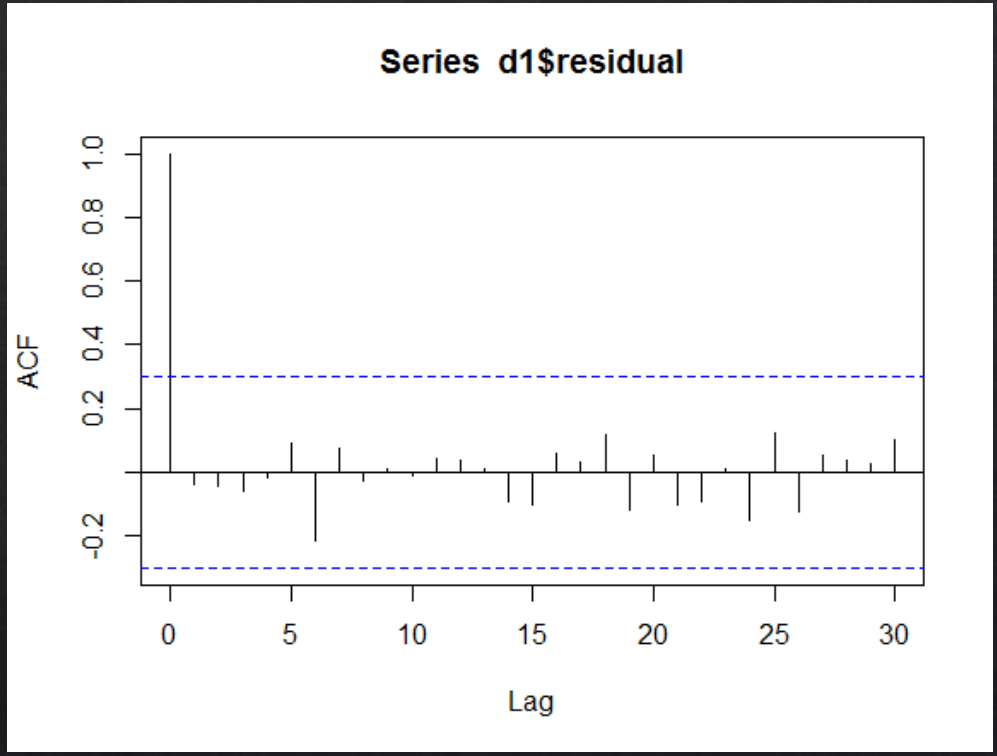
arima(x = x, order = c(2, 1, 0), seasonal = list(order = c(1, 1, 0), period = 4))

Coefficients:

	ar1	ar2	sar1
	-0.3894	-0.3291	-0.3348
s.e.	0.1591	0.1582	0.1681

sigma^2 estimated as 783.9: log likelihood = -176.17, aic = 360.34

ACF of the Residual



Excercise

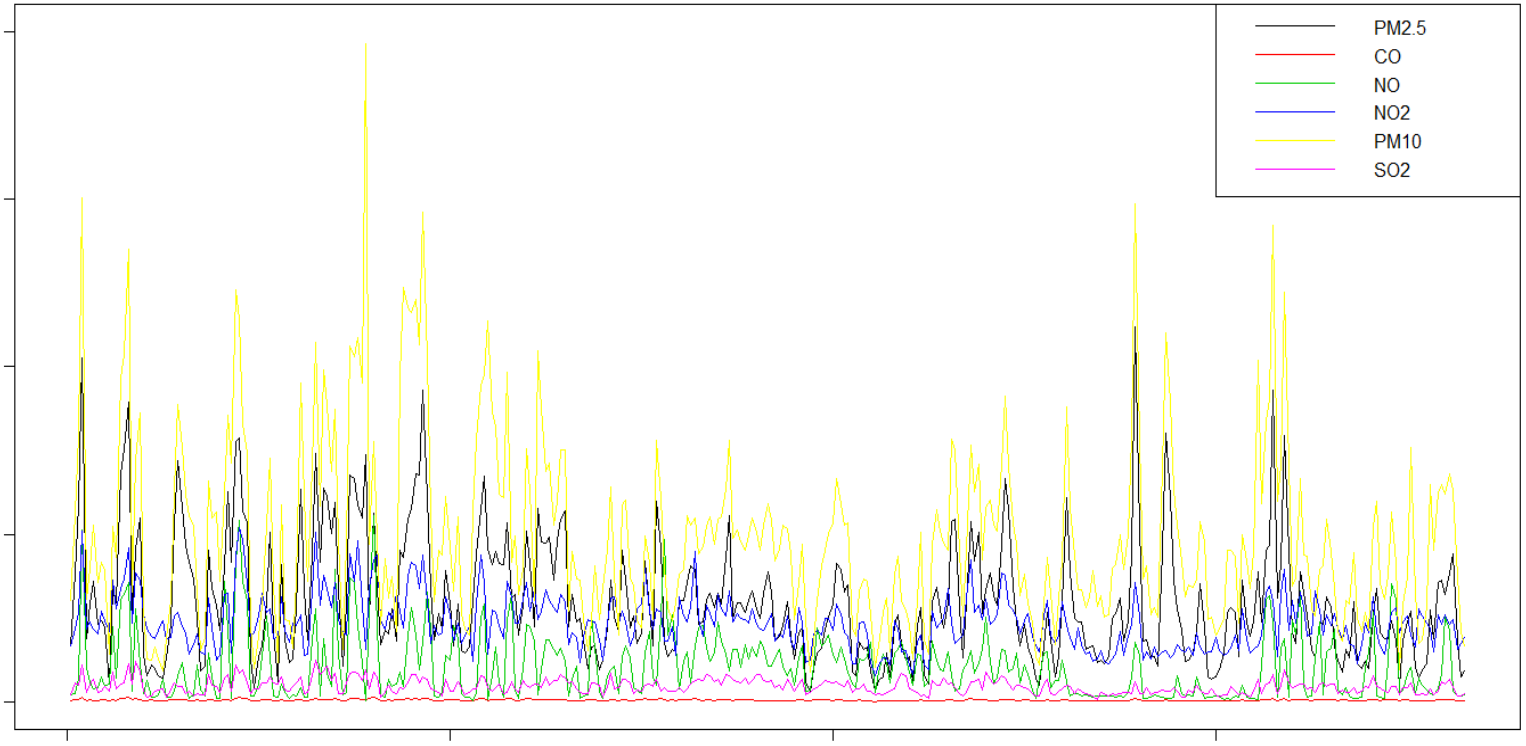
- Use **pe.csv** or rsp.csv to build a seasonal ARIMA model.
- You should show what model you use
  and how you get there as in the example.

空汙資料分析

（PM2.5）

資料介紹

- 資料來自於三義測站和陽明測站
  - 每小時測一次，並記錄24小時。三義共1822天，陽明共2043天。
- 各項汙染物時間趨勢圖(土城測站)



檔案格式說明

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	PM25	PM10	CO	NO	NO2	NOx	O3	PH_RAIN	RAINFALL	RH	SO2	TEMP	UVB	WIND_DIREC	WIND_SPEED
2	20	35	0.12	1.21	13.27	14.48	22.8	0	0	90.51	1.5	14.77	0	33.16	4.24
3	16	28	0.11	0.98	9.17	10.15	25.7	0	0	90.25	1.4	14.91	0	30.16	4.36
4	16	26	0.33	0.78	7.98	8.76	26.9	0	0	90.24	1.5	15.03	0	35.15	4.27
5	15	33	0.32	1.08	10.51	11.58	25.4	0	0	90.41	2.3	14.95	0	35.3	4.01
6	13	30	0.32	1.07	15.44	16.51	16.4	0	0	91.48	2.4	14.72	0	34.57	4.13
7	16	31	0.3	0.51	11.15	11.66	20.7	0	0	90.81	2.8	14.8	0	37.34	4.51
8	21	34	0.29	0.45	10.75	11.2	22.4	0	0	90.92	2.5	14.94	0	43.44	4.44
9	27	36	0.28	0.84	8.19	9.03	25	0	0	90.31	2	15.04	0.2	35.52	5.01
10	26	31	0.28	1.22	5.54	6.77	26.9	0	0	87.83	1.8	15.57	0.9	47.45	4.59
11	27	30	0.28	1.78	4.79	6.57	24.7	0	0	82.11	1.6	17.29	2.2	49.02	2.99
12	21	32	0.34	3.7	8.19	11.89	22.4	0	0	68.62	1.7	20.79	4	44.86	3.64
13	13	32	0.37	4.41	9.03	13.44	24.9	0	0	72.67	2.4	20.64	5.1	75.47	4.2
14	16	38	0.37	4.05	11.19	15.25	32.7	0	0	73.2	3.4	20.52	5.3	40.72	3.7
15	19	48	0.32	2.56	8.15	10.7	42.3	0	0	77.42	2.9	19.41	4.4	65.55	6.18
16	20	52	0.31	2.1	8.82	10.93	40.1	0	0	82.15	2.5	17.81	2.4	50.77	5.38
17	20	49	0.33	2.31	12.01	14.32	35.4	0	0	83.91	2.7	17.31	1	59.66	5.38
18	23	42	0.36	1.77	12.7	14.47	32.5	0	0	86.04	2.3	16.71	0.3	54.11	6.19
19	28	37	0.4	0.97	14.13	15.09	28.8	0	0	91.32	2.2	15.37	0	60.86	6.01
20	27	42	0.37	0.65	15.04	15.69	24.9	0	0	91.5	2.4	15.02	0	49.02	5.54
21	24	45	0.38	0.75	14.7	15.45	23.1	0	0	90.48	2.2	14.99	0	45.84	5.67
22	25	43	0.36	0.82	14.27	15.09	22.6	0	0	89.22	1.7	14.97	0	36.26	5.03
23	27	45	0.37	0.82	15.63	16.45	18.5	0	0	89.85	1.8	14.88	0	33.52	4.42
24	25	49	0.37	0.68	12.75	13.43	18.4	0	0	89.95	2	14.78	0	39.9	3.86
25	26	46	0.35	0.74	9.98	10.72	21	0	0	89.62	2.1	14.68	0	45.63	5.1
26	22	35	0.35	0.83	11.07	11.9	20.2	0	0	89.31	2	14.58	0	38.63	5.2
27	15	31	0.35	0.58	12.74	13.32	17.7	0	0	90.47	2.1	14.25	0	36.03	4.39
28	19	36	0.33	0.64	11.44	12.08	18.6	0	0	90.21	1.9	14.1	0	39.25	4.37
29	18	35	0.31	0.62	10.92	11.54	19.2	0	0	90.32	2.4	13.97	0	39.23	4.35

測目名稱說明(三義測站空污資料)

測項簡稱	單位	測項名稱
CO	ppm	一氧化碳
NMHC	ppm	非甲烷碳氫化合物
NO	ppb	一氧化氮
NO2	ppb	二氧化氮
Nox	ppb	氮氧化物
O3	ppb	臭氧
PH_RAIN	pH	酸鹼值(酸雨)
PM2.5	μg/m3	懸浮微粒
PM10	μg/m3	細懸浮微粒
RAIN_COND	μS/cm	導電度(酸雨)
RH	%	濕度
SO2	ppb	二氧化硫
TEMP	℃	溫度
UVB	UVI	紫外線指數
RAINFALL	mm	雨量
WIND_DIREC	風向(以每小時最後10分鐘向量平均)	風向
WIND_SPEED	風速(以每小時最後10分鐘算術平均)	風速

作業要求

1. 對PM2.5建立迴歸模型(Regression with time series residuals models)
   - Seasonal? Non-seasonal? AR? ARMA? ARIMA?
2. 估計參數和模型選擇。
   - 使用R來分析資料(可以上網找一些open source來用)
   - 說明估計方法(Nonlinear CLS? Iterative method?)
   - 解釋模型(繪出汙染物時間趨勢圖來解釋迴歸合理性。選出哪些重要參數？使用了何種時間序列的模型？模型是否吻合我們對PM2.5的了解？)
3. 預測兩天未來資料
   - 解釋方法(一步步做？一次做？)

評分標準

1. 程式
2. 估計方法
3. 解釋模型
4. 預測
5. 呈現方式

背景介紹(PM2.5)

- 懸浮微粒又稱氣膠(aerosol)、氣融膠或氣膠粒子
- 小於10微米的稱為懸浮微粒(PM10)，小於或等於2.5微米稱為細懸浮微粒(PM2.5)
- 「若吸入的氣膠大於 1 微米，會被呼吸道攔截下來；0.3~1 微米容易隨氣流方向移動，吸進後再被呼出去的機率高；小於 0.3 微米的氣膠，布朗運動效應明顯（粒子會做連續快速而不規則的隨機移動），氣膠容易碰撞到肺部組織而沾黏至潮濕的表面，留在體內影響健康。」

1.懸浮微粒--常被忽略的大氣成分 科學Online http://highscope.ch.ntu.edu.tw/wordpress/?p=23380
2.無所不在的氣膠 科學人 http://news.sina.com.tw/magazine/article/6187.html
3.環保署，空氣品質改善維護資訊網 http://air.epa.gov.tw/Public/suspended_particles.aspx

背景介紹(生成原因)

依來源區分		依性質區分	
自然界產出	人類行為產出	原生性PM2.5	衍生性PM2.5
火山爆發、地殼岩石等	石化燃料及工業排放、移動源廢氣等燃燒行為。	直接從自然與人為活動所排放，在大氣環境中未經化學反應的微粒－如天然的海鹽飛沫、營建工地粉塵、車行揚塵及工廠直接排放。	自然與人為活動排放到大氣環境中的化學物質經過太陽光照或其他化學反應後生成－如燃煤、燃油及燃氣電廠、煉鋼廠、石化相關產業工廠、機動車輛、船舶、建物塗料、農業施肥、禽畜排泄及生活污水等。

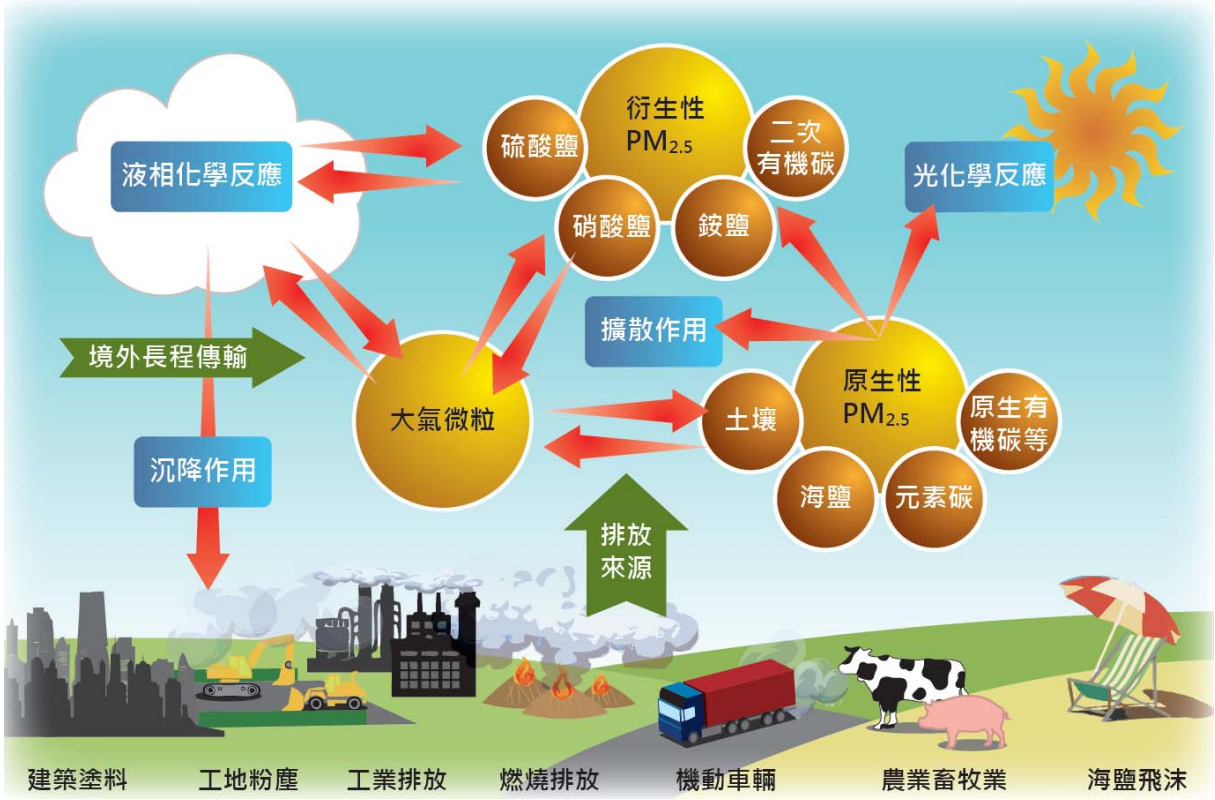
出處：環保署，空氣品質改善維護資訊網 http://air.epa.gov.tw/Public/suspended_particles.aspx

背景介紹(來源)

- 國內產生
  - 工業
  - 機動車輛
  - 營建工程與農業
  - 其它：如民生活動、燒金紙
- 境外輸入
  - 中國霾害（冬、春季）
  - 東南亞生質燃燒（３～４月）
  - 東北亞工業污染源

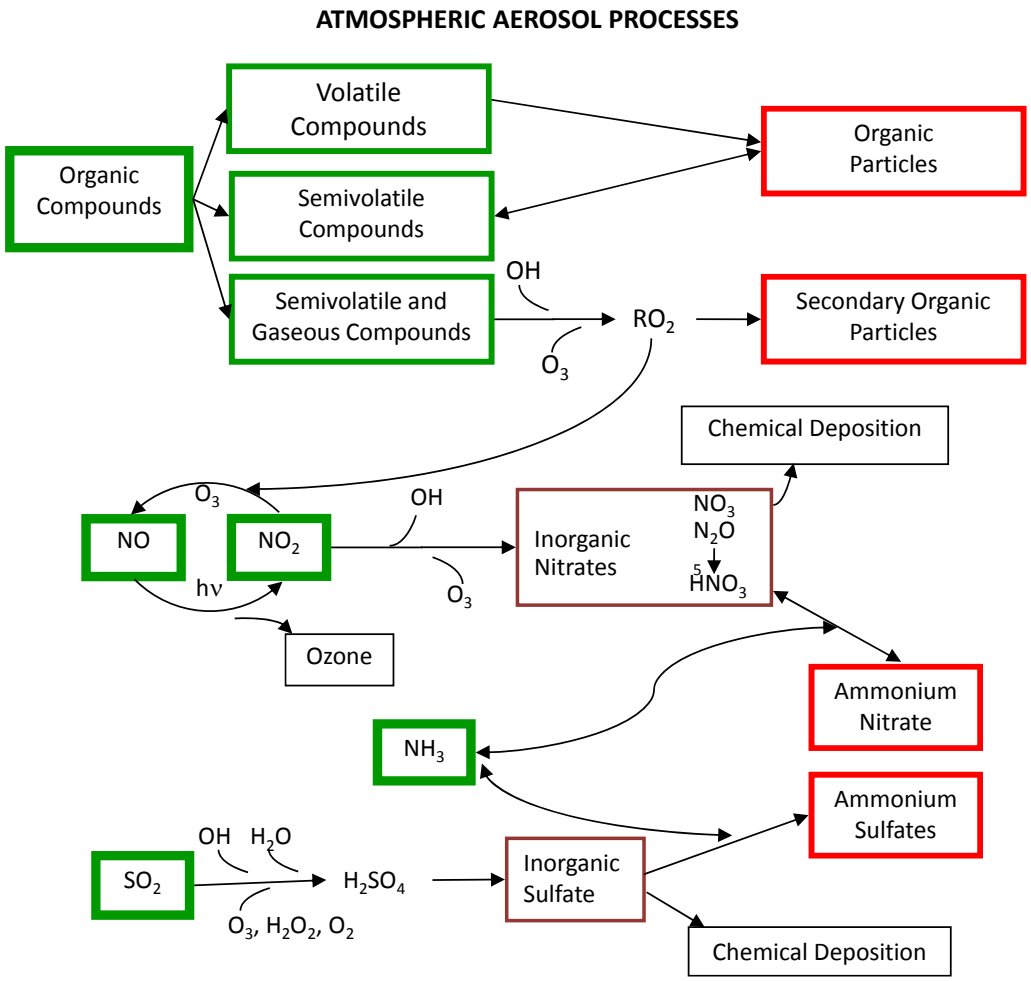
出處：環保署，空氣品質改善維護資訊網 http://air.epa.gov.tw/Public/suspended_particles.aspx

細懸浮微粒反應生成機制



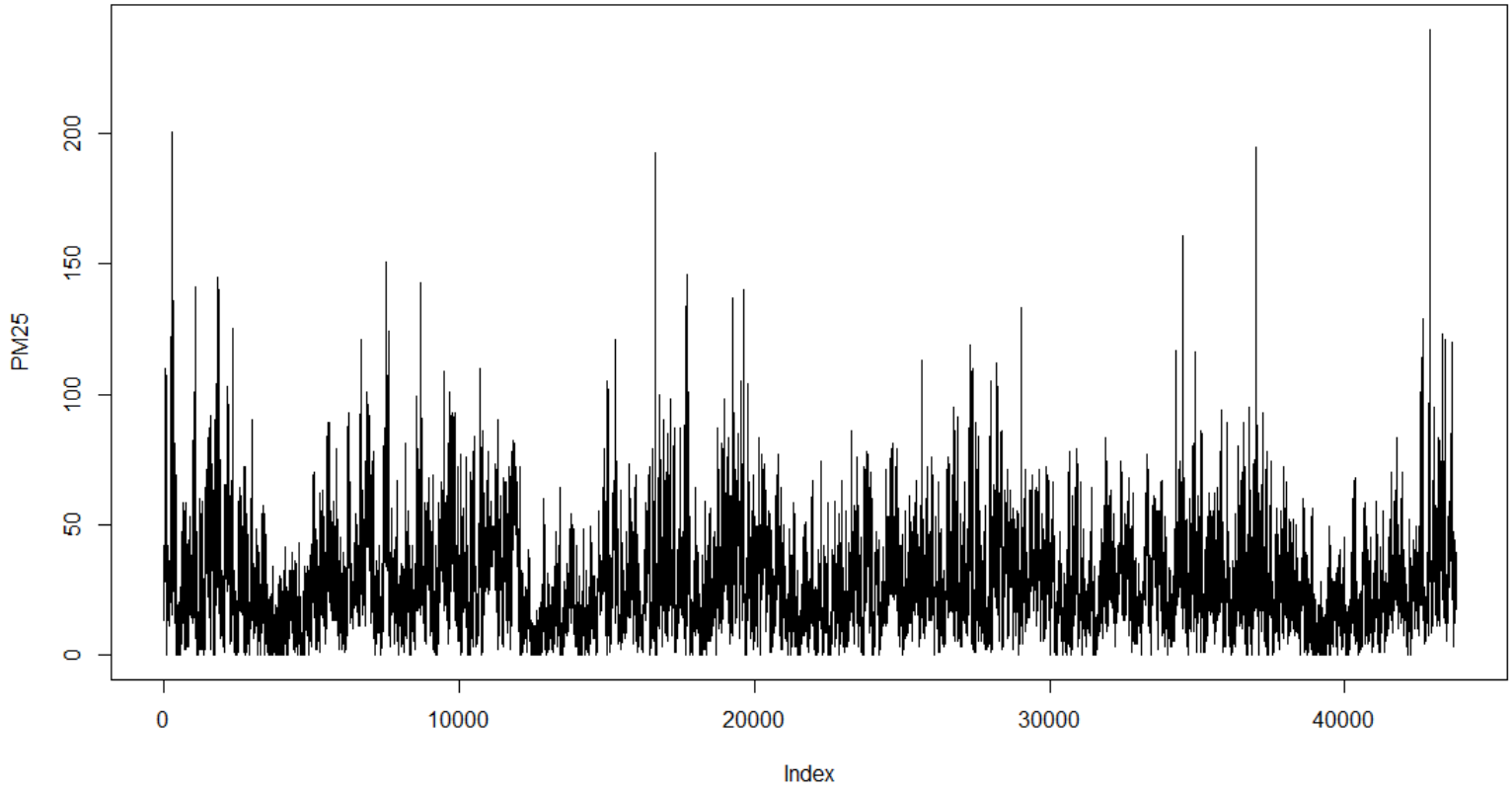
出處：環保署，空氣品質改善維護資訊網 http://air.epa.gov.tw/Public/suspended_particles.aspx

細懸浮微粒反應生成機制

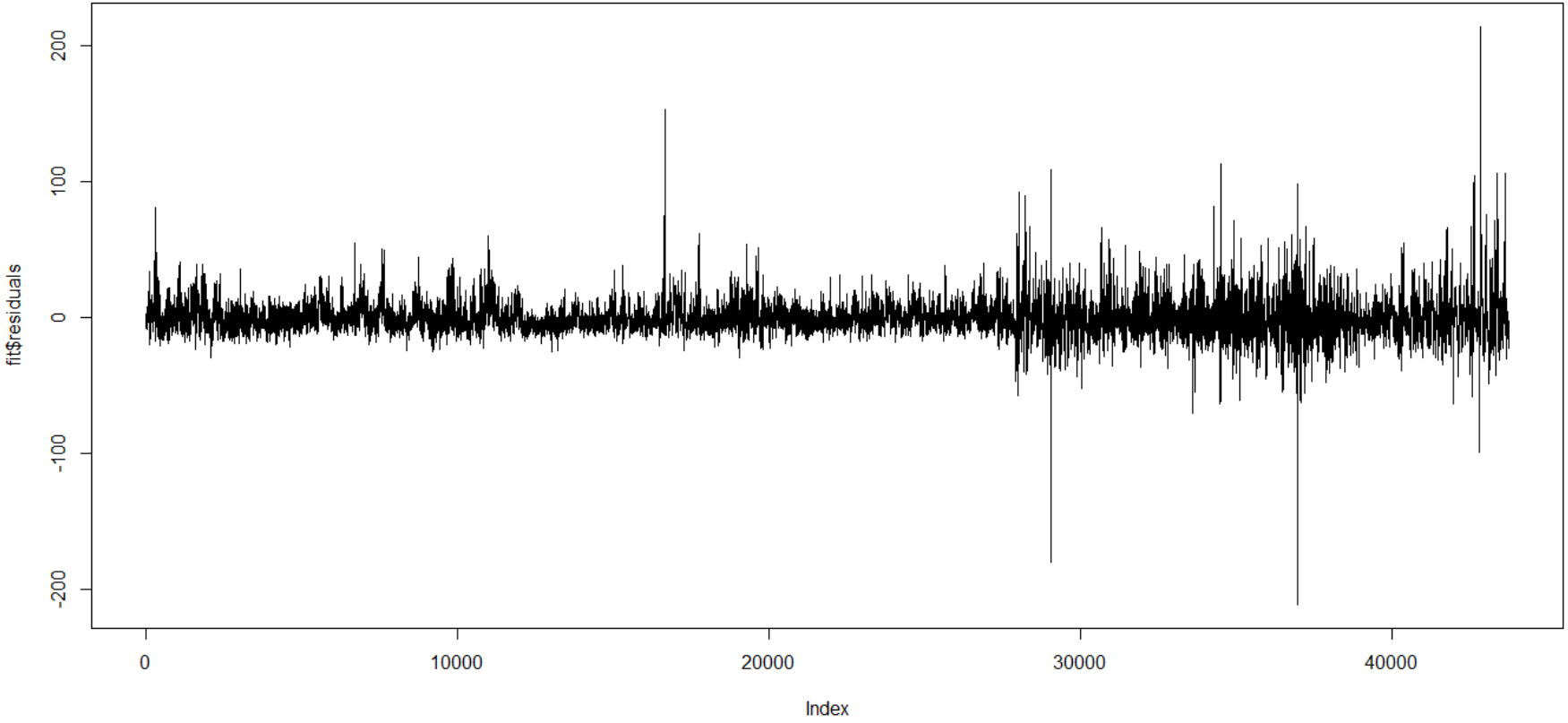


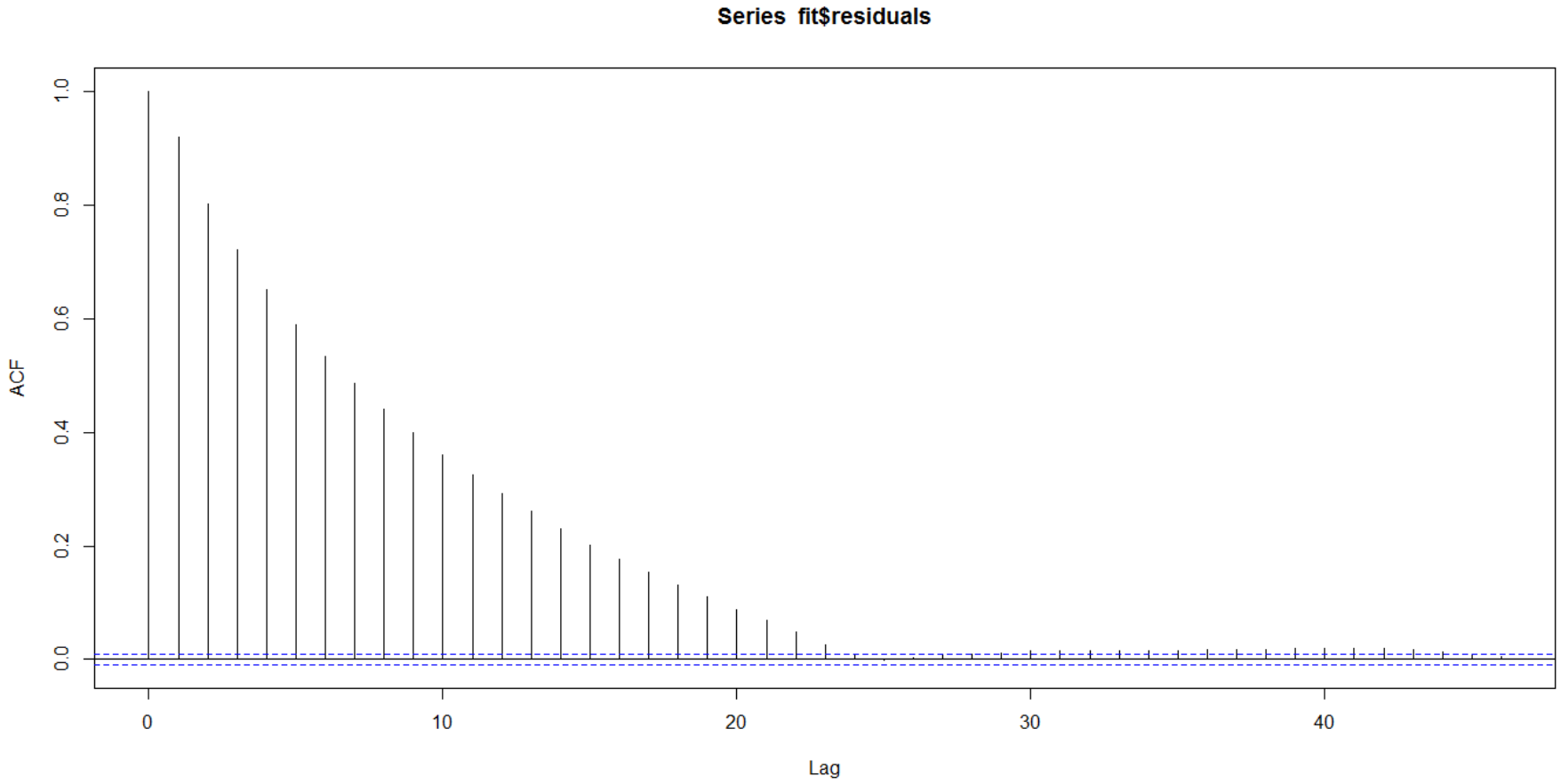
出處：環保署，空氣品質改善維護資訊網

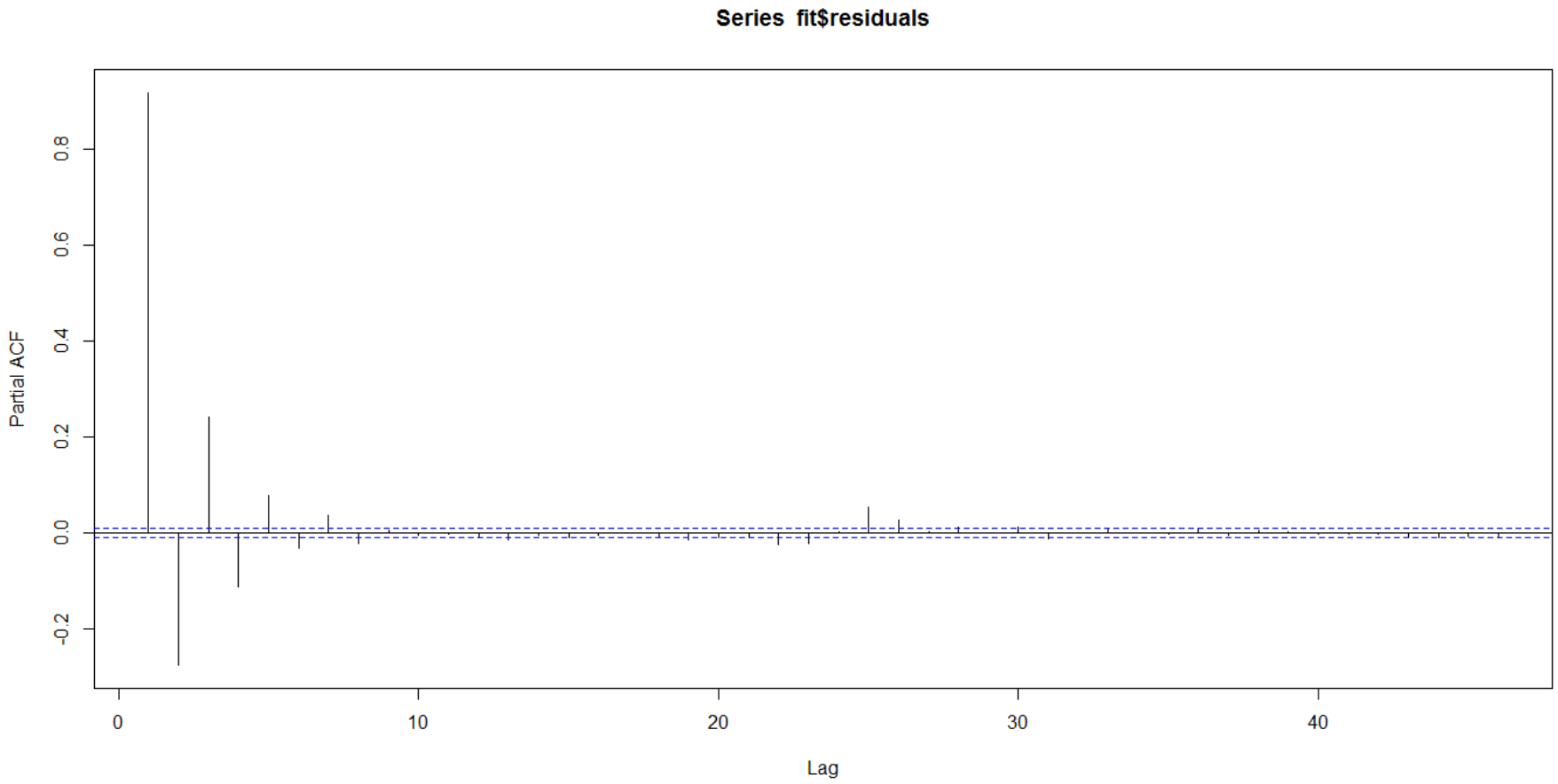
PM2.5

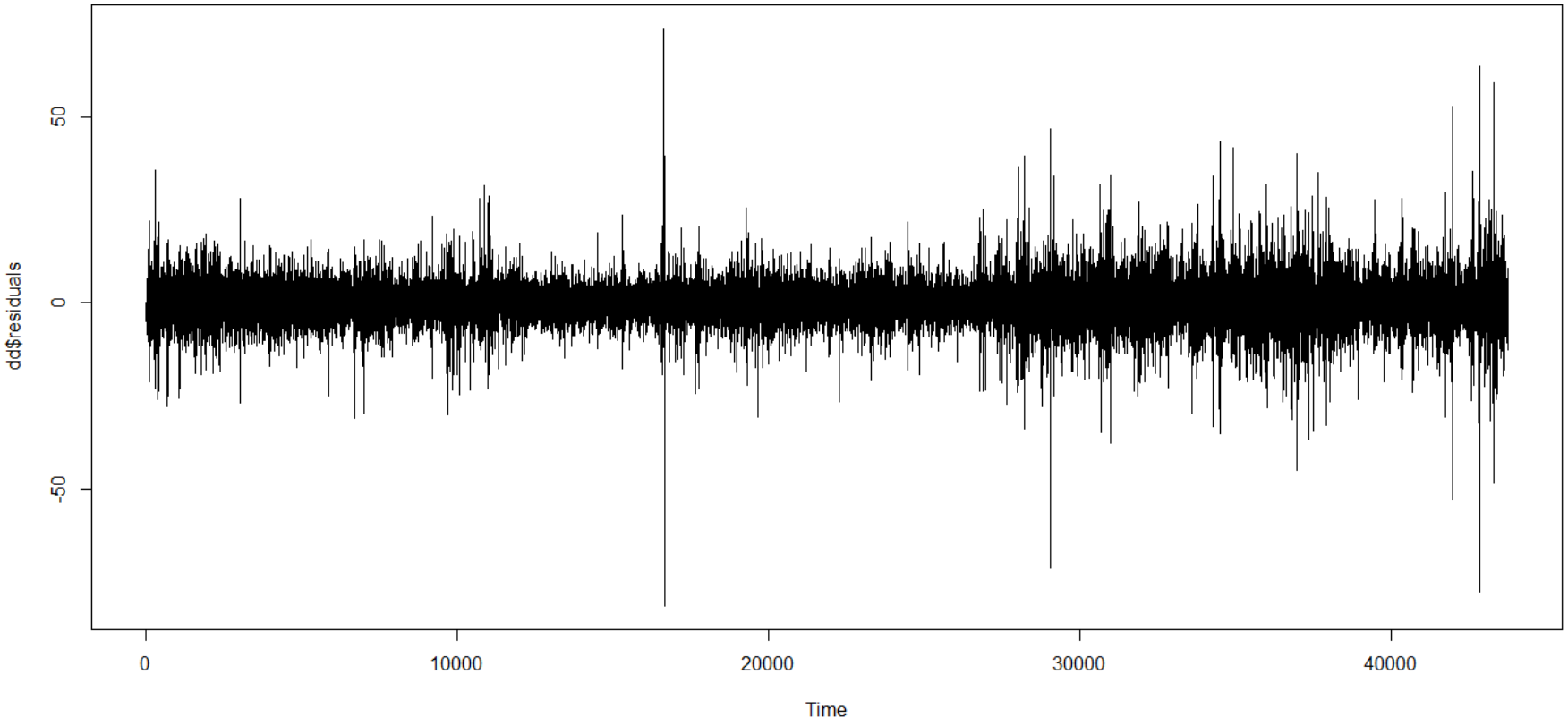


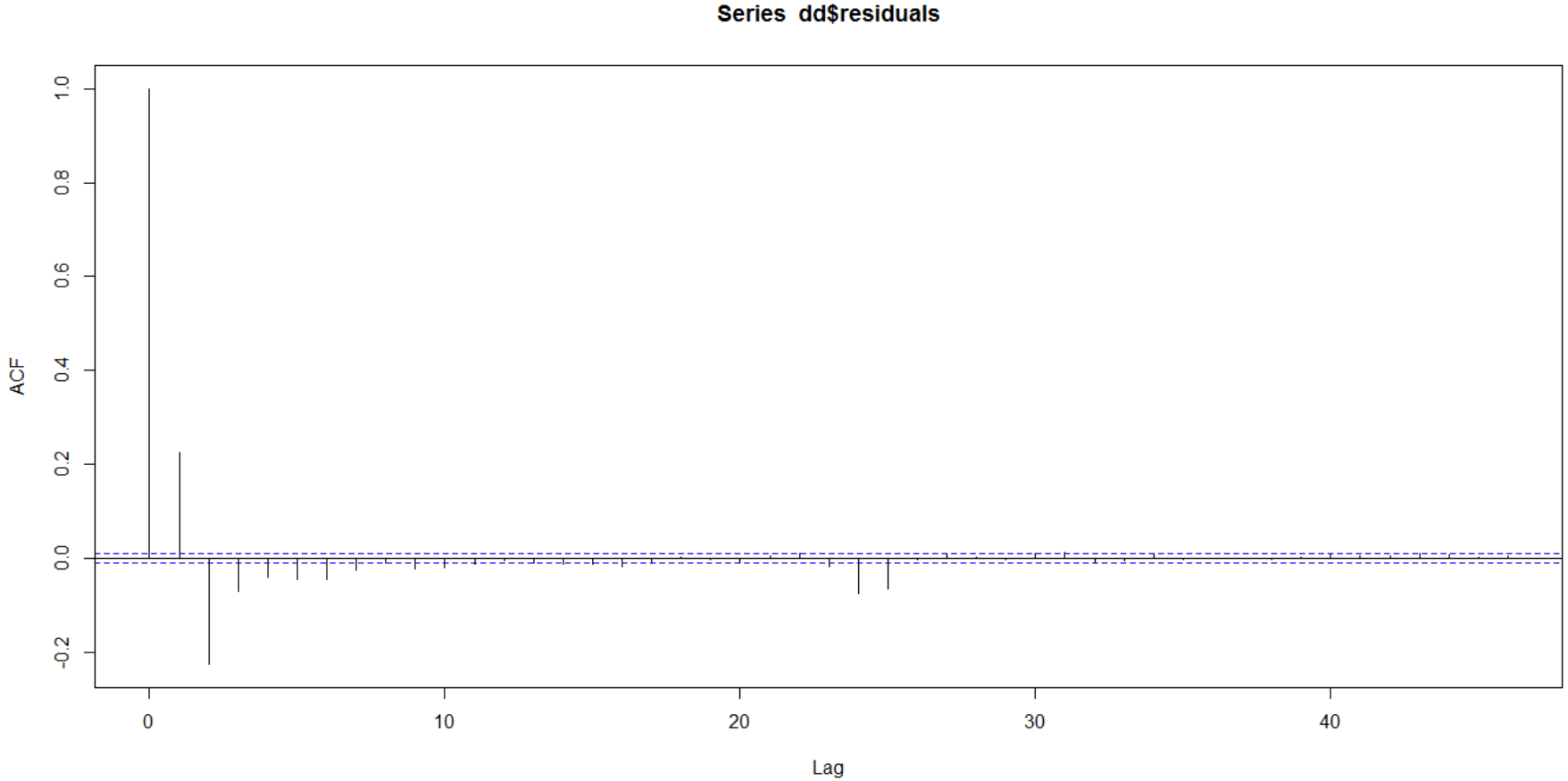
做完迴歸的殘差

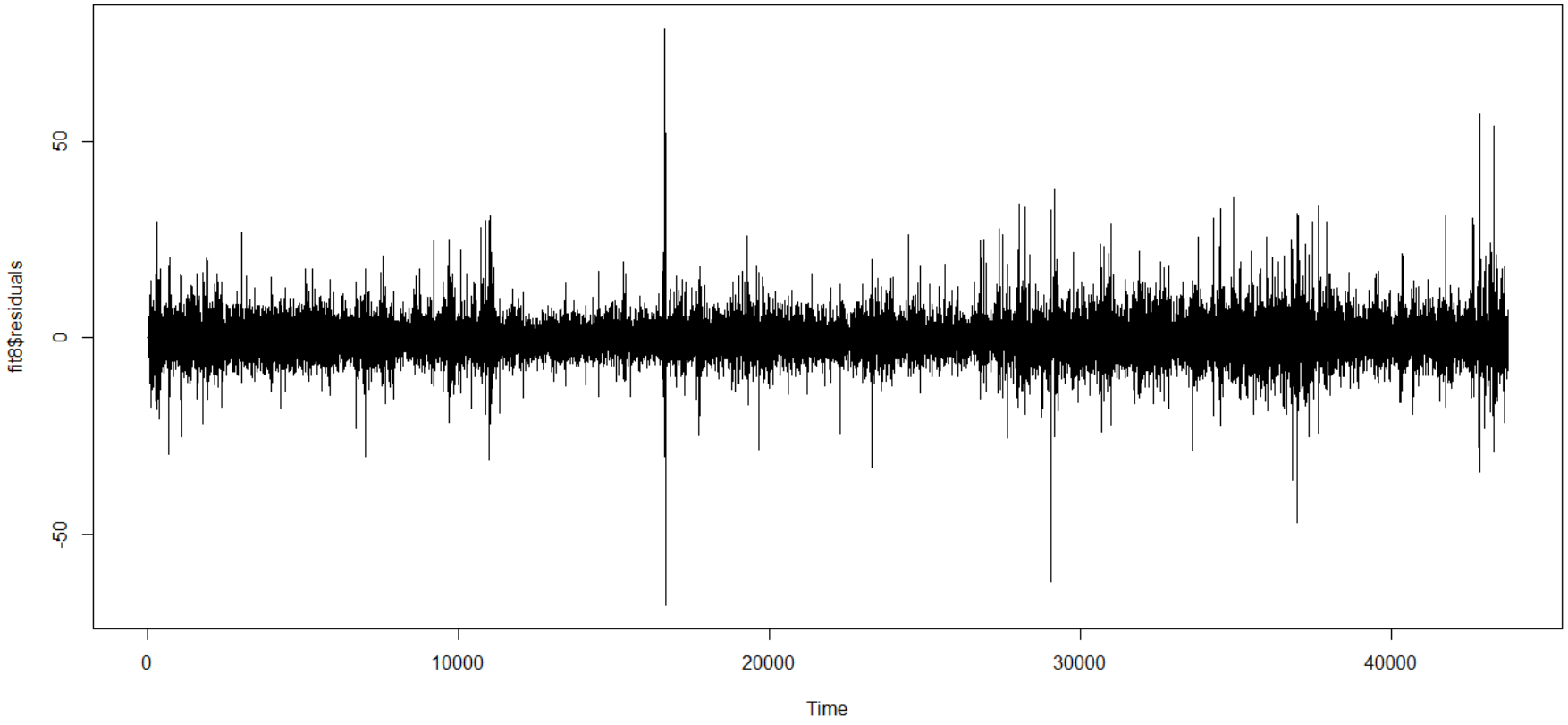

fit$residuals
200
100
0
-100
-200
0
10000
20000
30000
40000
Index

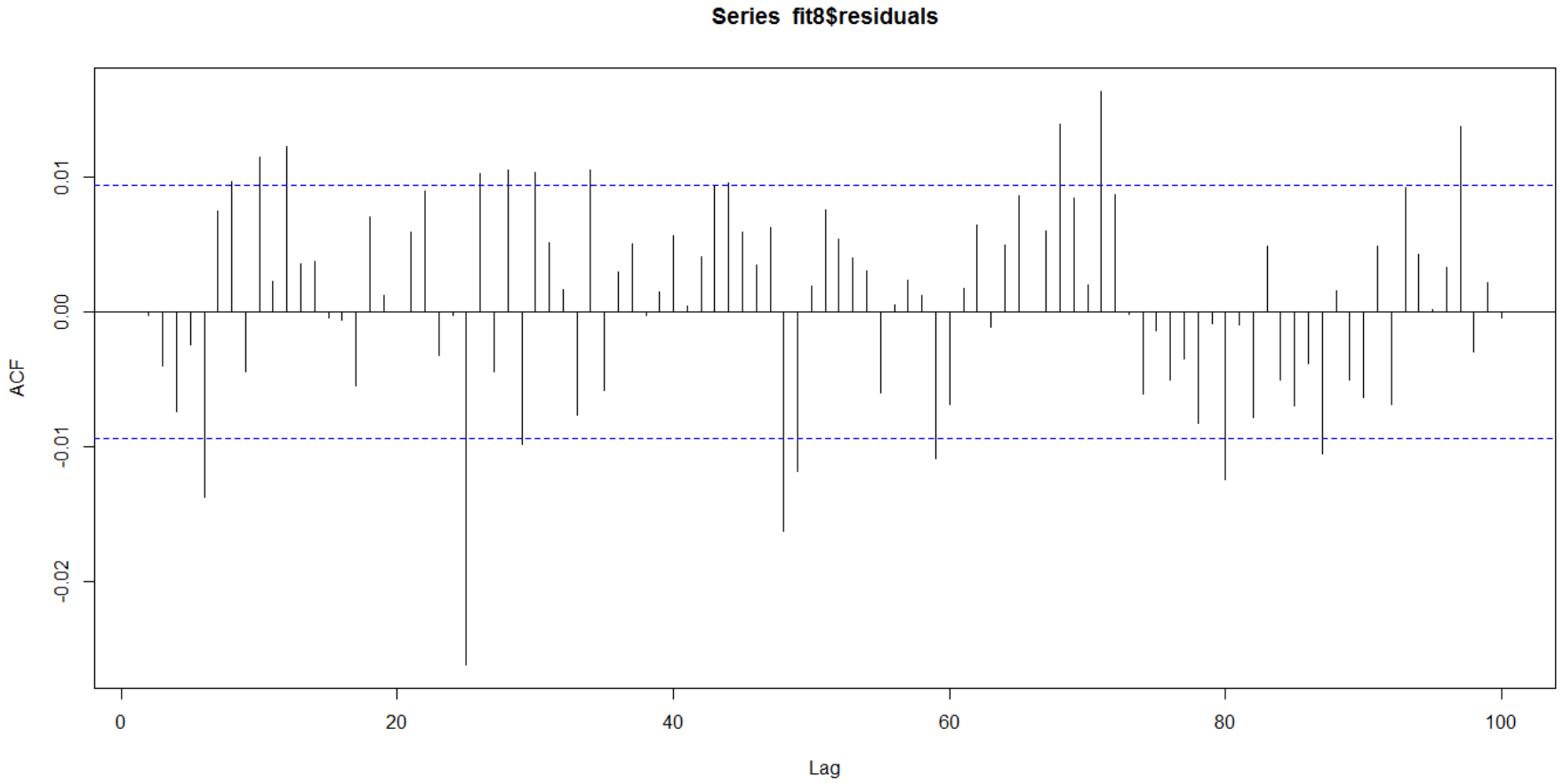

Series fit$residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Lag

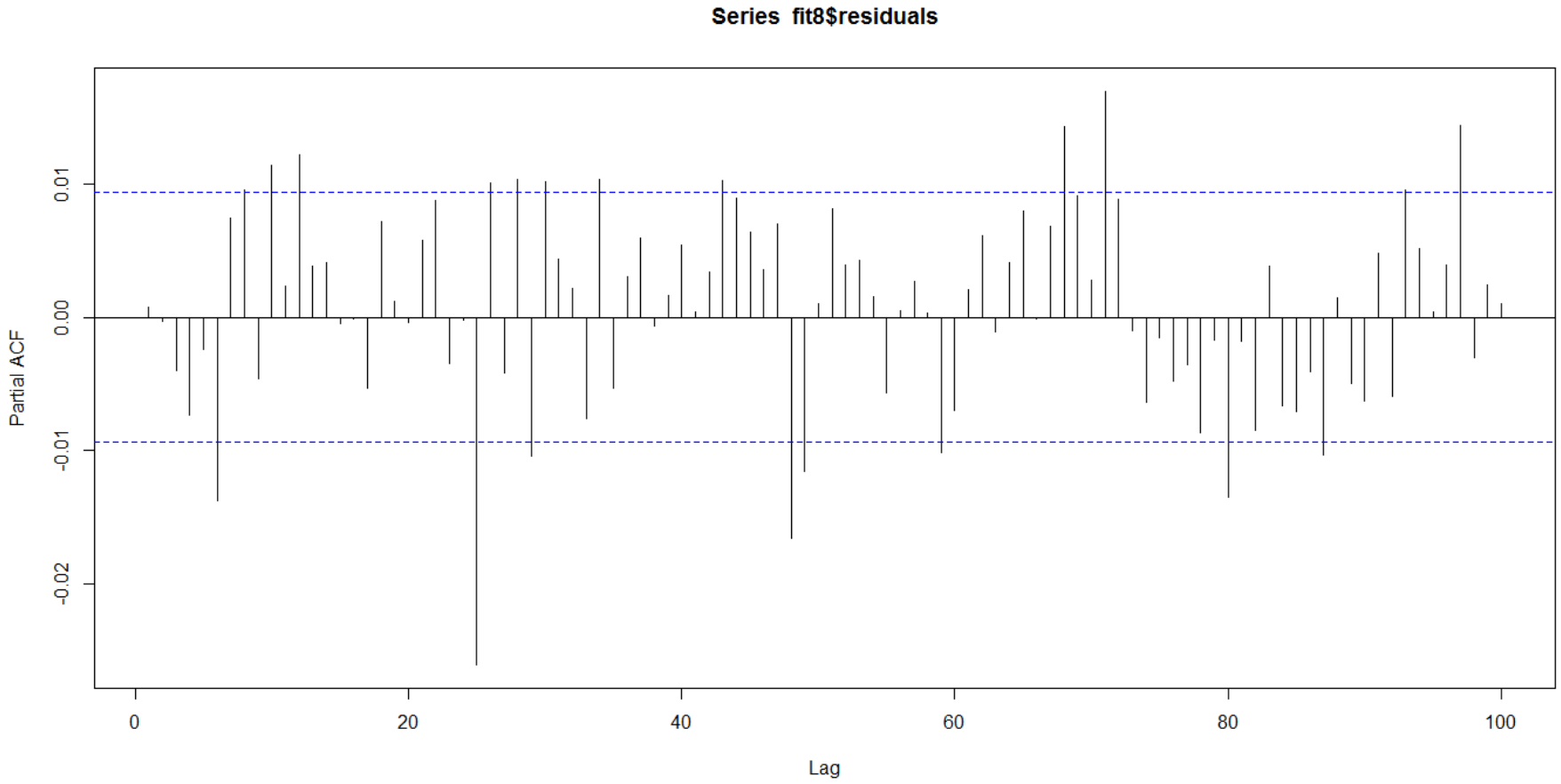

Series fit$residuals
Partial ACF
0.8
0.6
0.4
0.2
0.0
-0.2
0
10
20
30
40
Lag


dd$residuals
50
0
-50
0
10000
20000
30000
40000
Time


Series dd$residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
10
20
30
40
Lag


fit8$residuals
50
0
-50
0
10000
20000
30000
40000
Time


Series fit8$residuals
ACF
0.01
0.00
-0.01
-0.02
0
20
40
60
80
100
Lag


Series fit8$residuals
Partial ACF
0.01
0.00
-0.01
-0.02
0
20
40
60
80
100
Lag

時間序列期末報告
財務資料

資料介紹

- 政府一年期與三年期公債殖利率
- 公債殖利率可視為市場認為政府借貸的利率水準

資料介紹

y1	y3
3.24	3.7
3.32	3.75
3.29	3.8
3.26	3.77
3.29	3.8
3.29	3.76
3.31	3.71
3.29	3.65
3.2	3.52
3.15	3.52
3.1	3.52
2.99	3.37

y1 一年期公債殖利率
y3三年期公債殖利率

由舊到新

期末報告要求

- 預測接下來五期 一年殖利率與三年殖利率
  1. 利用時間序列的模型
     - 建構預測的模型、參數估計結果、選模結果
  2. 利用一年期與三年期殖利率的相關性與時間序列模型
     - 建構模型說明一年與三年期殖利率的關係，以及是否是否會受到過去歷史的影響，如果是 受落後幾期影響
     - 建構預測的模型、參數估計結果、選模結果

- 預測結果越接近真實值分數越高!!!