Question 1

> yld=yield
> length(yld)

[1] 156

> yldfit=arima(yld, order =c(1,0,0))
> yldfit

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

Coefficients:
    ar1 intercept
     0.8517    0.9666
   s.e.  0.0424   0.0820

sigma^2 estimated as 0.02447:  log likelihood = 67.41,  aic = -128.81

> Question 1(a).

> tsdiag(yldfit)
>
> Box.test(yldfit$res, lag = 30, fitdf=1, type="Ljung-Box")

Box-Ljung test

data: yldfit$res
X-squared = 32.4372, df = 29, p-value = 0.301

> AR(1) choice of model for original data yield justified by residual plot, sample acfs of fitted residuals, and Box-Ljung test. (P-values are high.)

Question 1(b).

> yldfit1=arima(yld, order=c(2,0,0))
> yldfit1
Call:
arima(x = yld, order = c(2, 0, 0))

Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>ar1</th>
<th>ar2</th>
<th>intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.9500</td>
<td>-0.1140</td>
<td>0.9704</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.0839</td>
<td>0.0842</td>
<td>0.0742</td>
</tr>
</tbody>
</table>

sigma^2 estimated as 0.02418: log likelihood = 68.32, aic = -128.64

> yldfit2=arima(yld, order=c(1,0,1))
> yldfit2

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

Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>ar1</th>
<th>ma1</th>
<th>intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.8097</td>
<td>0.1583</td>
<td>0.9707</td>
</tr>
<tr>
<td>s.e.</td>
<td>0.0567</td>
<td>0.1010</td>
<td>0.0740</td>
</tr>
</tbody>
</table>

sigma^2 estimated as 0.02411: log likelihood = 68.52, aic = -129.05

> 

Standardized estimated overfitting coefficient ratios are not large; and AIC’s of overfitted models are comparable or larger than that of original model.

**Question 1(c).**

> predict(yldfit, n.ahead=3, se.fit=T)

$pred
Time Series:
Start = 157
End = 159
Frequency = 1
[1] 0.5606803 0.6208791 0.6721509

$se
Time Series:
Start = 157
End = 159
Frequency = 1
[1] 0.1564261 0.2054729 0.2347233

> p=predict(yldfit, n.ahead=3, se.fit=T)
> z=qnorm(.99)
> c(p$pred[1:3] - z*p$se[1:3],p$pred[1:3] + z*p$se[1:3])

[1] 0.1967787 0.1428776 0.1261028 0.9245819 1.0988806 1.2181989

The CI’s, respectively, are (0.1967787, 0.9245819), (0.1428776, 1.0988806), (0.1261028, 1.2181989).

Question 2.

humid.week

> hum=humid.week
> length(hum)

[1] 195

> par(mfrow = c(3,1))
> ts.plot(hum)
> acf(hum)
> acf(hum, type= "partial")
>
 Seems to be non-stationary: due to trend up? Try differencing once. We would expect a seasonal variation with periodicity 52 since data is weekly, so keep lag 52 differencing in mind if fit with first difference no good.

```r
> hum3=diff(hum)
>
> par(mfrow=c(3,1))
> ts.plot(hum3)
> acf(hum3)
> acf(hum3, type="partial")
> ```
Try MA(1)

> hum3fit=arima(hum3,order=c(0,0,1))
> hum3fit

Call:
arima(x = hum3, order = c(0, 0, 1))

Coefficients:
    ma1  intercept
           -0.6792  0.0701
s.e.  0.0534  0.1518

sigma^2 estimated as 42.49:  log likelihood = -639.26,  aic = 1284.51

>
> tsdiag(hum3fit)

> Box.test(hum3fit$res, lag = 20, type="Ljung-Box", fitdf=1)

Box-Ljung test

data:  hum3fit$res
X-squared = 17.6298, df = 19, p-value = 0.5473

> MA(1) fit on differenced data justified by residual plot, acf plot of residuals, and Ljung-Box test.

nitrogen
Daily readings tsplot reveals a high spike about mid-February. Looks like AR(1) (geometric decline of acf, one large spike at lag 1 in pacf.)

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

Coefficients:
    ar1  intercept
    0.7138  2.5706
s.e.  0.0525  0.1493

sigma^2 estimated as 0.3377: log likelihood = -158.05, aic = 322.11

> 

> tsdiag(nitfit)

> 

> Box.test(nitfit$res, lag=20, type = "Ljung-Box", fitdf=1)
Box-Ljung test

data: nitfit$res
X-squared = 8.0089, df = 19, p-value = 0.9866
>
AR(1) fit on *original* data justified by residual plot and Ljung-Box test.