STOR 356 FINAL EXAM
MAY 1 2008

This is an open book exam. Course text, personal notes and calculator are permitted. You have 3 hours to complete the test. Personal computers are not allowed. If you have any queries about the meaning of the questions, ask the instructor for assistance.

Answer all questions. You are expected to show your full working, but any results that you take from the text or from course notes you may quote without giving any derivation of them.

Where the question asks for a numerical answer, I will always accept a formula (for example, $(0.6)^2 - 5.7\sqrt{12.333}$), so long as it's an *explicit* expression that I could verify numerically. However, the actual numerical answer (in the above case, –19.66) will always be accepted so long as it's clear how it was obtained.

Important Note. In cases where a later part of a question relies on the answer to an earlier part, an incorrect answer to the earlier part will not prevent you obtaining full credit for the later part, if the method used in the later part is correct.

1. (**50 points total.**) Consider the process

$$X_t - 0.2X_{t-1} + 0.48X_{t-2} \quad = \quad Z_t - 0.7Z_{t-1}$$

where $Z_t \sim WN[0, \sigma^2]$, $\sigma^2 = 3$.

(a) (**6 points.**) Is the process (i) stationary, (ii) causal, (iii) invertible? In each case provide a short explanation for your answer.

(b) (**9 points.**) Consider the ψ expansion of the process, $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$. Show that $\psi_1 = -0.5$, $\psi_3 = 0.124$, and find ψ_2.

(c) (**9 points.**) Consider the π expansion of the process, $Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$. Show that $\pi_1 = 0.5$, $\pi_3 = 0.581$, and find π_2.

If you didn't succeed in solving (b) and (c), assume the results $\psi_2 = -0.5$, $\pi_2 = 0.6$ for the remaining parts of the question (these are not the correct answers). Also assume $\pi_4 = .407$ (which is the correct answer but you are not asked to prove that).

(d) (**8 points.**) What are the optimal predictors for the one-step and two step predictors, (i) $\tilde{P}_t X_{t+1}$, (ii) $\tilde{P}_t X_{t+2}$?

[The answer will be some linear combination of X_s, $s \le t$. Give the explicit numerical coefficients for X_t, X_{t-1} and X_{t-2}.]

(e) (**8 points.**) What are the mean squared prediction errors of $\tilde{P}_t X_{t+1}$, $\tilde{P}_t X_{t+2}$ and $\tilde{P}_t X_{t+3}$?

(f) (**10 points.**) Assume we want to calculate the theoretical values of $\rho_X(h)$ for $h = 1, 2, 3$. Write down a set of equations that these values must satisfy. (You are not expected to obtain actual numerical values.)

2. (**25 points total.**) Figure 1 (page 8, top plot) is a plot of part of the data file chocs.tsm, described as “Chocolate-based confectionery production. Monthly production in tonnes.” This portion of the series is from January 1959 to December 1974 ($N = 192$). We work with the logarithms of chocolate sales and their differences. Differences of lag 12 are shown in the middle plot of Figure 1; differences of lag 12 followed by differences of lag 1 are shown in the bottom plot of Figure 1. The sample ACF and PACF of all three series are shown in Figure 2.

A linear regression was fitted by OLS (see top plot in Figure 1) and an ARMA model was fitted to the residuals using the `autofit` function with max p and q set to 15. The results were as follows:

```
ARMA Model:
X(t) = - .06553 X(t-1) - .1283 X(t-2) - .06600 X(t-3) - .1183 X(t-4)
       - .05560 X(t-5) - .1284 X(t-6) - .08211 X(t-7) - .1241 X(t-8)
       - .06677 X(t-9) - .1174 X(t-10) - .06707 X(t-11) + .8643 X(t-12)
     + Z(t) + .3281 Z(t-1) + .4294 Z(t-2) + .2283 Z(t-3)
     + .1772 Z(t-4) + .2500 Z(t-5) + .2851 Z(t-6) + .1945 Z(t-7)
     + .3547 Z(t-8) + .09591 Z(t-9) + .2048 Z(t-10) + .2743 Z(t-11)
     - .4471 Z(t-12)
WN Variance = .007680

AR Coefficients
      -.065530      -.128341      -.066003      -.118297
      -.055602      -.128440      -.082107      -.124098
      -.066770      -.117435      -.067068       .864257
Standard Error of AR Coefficients
       .071101       .070969       .071397       .071428
       .070399       .069922       .070594       .070822
       .071140       .071013       .070306       .069943

MA Coefficients
       .328126       .429356       .228260       .177203
       .249969       .285127       .194522       .354738
       .095913       .204782       .274261      -.447124
Standard Error of MA Coefficients
       .107898       .110199       .123145       .125534
       .113026       .113895       .112701       .110390
       .119186       .114761       .102373       .100635

(Residual SS)/N = .00767980

AICC = -.300530E+03
AICC = -.295142E+03   (Corrected for regression)
BIC  = -.261086E+03

-2Log(Likelihood) = -.358361E+03
Accuracy parameter = .100000E-08
```

```
Number of iterations = 6
Number of function evaluations = 496950
Uncertain minimum.
```

The results of a residual test were:

```
Ljung - Box statistic = 47.109 Chi-Square ( 20 ), p-value = .00057
McLeod - Li statistic = 32.160 Chi-Square ( 44 ), p-value = .90732
# Turning points = .13000E+03~AN(.12667E+03,sd = 5.8147), p-value = .56647
# Diff sign points = 93.000~AN(95.500,sd = 4.0104), p-value = .53304
Rank test statistic = .91360E+04~AN(.91680E+04,sd = .44512E+03), p-value = .94269
Jarque-Bera test statistic (for normality) = 1.3249 Chi-Square (2), p-value = .51559
Order of Min AICC YW Model for Residuals = 0
```

Forecasts for 10 steps ahead (on the scale of the original data, i.e. accounting for the logarithmic transformation and the linear trend) were

		Approximate 95 Percent Prediction Bounds	
Step	Prediction	Lower	Upper
1	.24677E+04	.20766E+04	.29323E+04
2	.43609E+04	.36492E+04	.52113E+04
3	.53886E+04	.44809E+04	.64803E+04
4	.48433E+04	.40229E+04	.58308E+04
5	.64001E+04	.53161E+04	.77051E+04
6	.54219E+04	.44979E+04	.65358E+04
7	.65470E+04	.54269E+04	.78982E+04
8	.63558E+04	.52683E+04	.76679E+04
9	.53839E+04	.44558E+04	.65052E+04
10	.50798E+04	.42026E+04	.61401E+04

A second analysis was performed based on the differenced series of order 12 (i.e. $\nabla_{12} Y_t$, where Y_t are the logged sales) and without any regression component. This time the results of `autofit` were:

```
ARMA Model:
X(t) = Z(t) + .2038 Z(t-1) + .3062 Z(t-2) + .1414 Z(t-3)
     + .08382 Z(t-4) + .1120 Z(t-5) + .05260 Z(t-6) + .007520 Z(t-7)
     + .2615 Z(t-8) + .08107 Z(t-9) + .1899 Z(t-10) + .1618 Z(t-11)
     - .6154 Z(t-12) - .1306 Z(t-13)
WN Variance = .007653
```

```
MA Coefficients
       .203797       .306153       .141392       .083816
       .111992       .052597       .007520       .261496
       .081068       .189921       .161765      -.615409
      -.130636
Standard Error of MA Coefficients
       .073897       .059895       .062951       .062237
       .062257       .059714       .059840       .059714
       .062257       .062237       .062951       .059895
       .073897
(Residual SS)/N = .00765281

AICC = -.319958E+03
BIC  = -.325207E+03
-2Log(Likelihood) = -.350504E+03
Accuracy parameter = .100000E-08
Number of iterations = 1
Number of function evaluations = 1871467
Uncertain minimum.
```

The results of the residual tests were:

```
Ljung - Box statistic = 24.745 Chi-Square ( 20 ), p-value = .21137
McLeod - Li statistic = 23.807 Chi-Square ( 33 ), p-value = .87973
# Turning points = .12800E+03~AN(.11867E+03,sd = 5.6283), p-value = .09726
# Diff sign points = 89.000~AN(89.500,sd = 3.8837), p-value = .89756
Rank test statistic = .78780E+04~AN(.80550E+04,sd = .40415E+03), p-value = .66142
Jarque-Bera test statistic (for normality) = 2.5075 Chi-Square (2), p-value = .28544
Order of Min AICC YW Model for Residuals = 0
```

The results for forecasting 10 steps ahead (on the original scale) were:

		Approximate 95 Percent Prediction Bounds	
Step	Prediction	Lower	Upper
1	.23678E+04	.19947E+04	.28107E+04
2	.42792E+04	.35923E+04	.50975E+04
3	.53738E+04	.44766E+04	.64510E+04
4	.49999E+04	.41584E+04	.60118E+04
5	.67337E+04	.55973E+04	.81009E+04
6	.56337E+04	.46782E+04	.67843E+04
7	.66412E+04	.55136E+04	.79993E+04
8	.60235E+04	.50008E+04	.72553E+04
9	.54551E+04	.45049E+04	.66058E+04
10	.52382E+04	.43235E+04	.63463E+04

A third analysis was performed based on the differenced series of order 12 and order 1 ($\nabla_1 \nabla_{12} Y_t$). This time the results of `autofit` were:

```
ARMA Model:
X(t) = Z(t) - .7728 Z(t-1) + .1401 Z(t-2) - .06174 Z(t-3)
     - .1242 Z(t-4) + .05444 Z(t-5) + .0009616 Z(t-6) - .1276 Z(t-7)
     + .3277 Z(t-8) - .1853 Z(t-9) + .08341 Z(t-10) + .07501 Z(t-11)
     - .8877 Z(t-12) + .6064 Z(t-13)
WN Variance = .008105

MA Coefficients
      -.772824       .140112      -.061736      -.124247
       .054440       .000962      -.127626       .327666
      -.185347       .083411       .075012      -.887682
       .606432
Standard Error of MA Coefficients
       .059431       .049665       .050446       .050272
       .049209       .042876       .041801       .042876
       .049209       .050272       .050446       .049665
       .059431
(Residual SS)/N = .00810508

AICC = -.303645E+03
BIC  = -.305287E+03
-2Log(Likelihood) = -.334206E+03
Accuracy parameter = .100000E-08
Number of iterations = 1
Number of function evaluations = 6021531
Uncertain minimum.
```

The results of the residual tests were:

```
Ljung - Box statistic = 94.645 Chi-Square ( 20 ), p-value = .00000
McLeod - Li statistic = 33.733 Chi-Square ( 33 ), p-value = .43189
# Turning points = .13100E+03~AN(.11800E+03,sd = 5.6125), p-value = .02054
# Diff sign points = 87.000~AN(89.000,sd = 3.8730), p-value = .60558
Rank test statistic = .79770E+04~AN(.79655E+04,sd = .40080E+03), p-value = .97711
Jarque-Bera test statistic (for normality) = 1.7251 Chi-Square (2), p-value = .42208
Order of Min AICC YW Model for Residuals = 11
```

The forecasts 10 steps ahead (on the original scale) were:

		Approximate 95 Percent Prediction Bounds	
Step	Prediction	Lower	Upper
1	.24051E+04	.20161E+04	.28692E+04
2	.44388E+04	.37041E+04	.53192E+04
3	.55736E+04	.45990E+04	.67547E+04
4	.52098E+04	.42670E+04	.63609E+04
5	.70666E+04	.57731E+04	.86498E+04
6	.60198E+04	.48972E+04	.73999E+04
7	.69769E+04	.56521E+04	.86123E+04
8	.65471E+04	.52992E+04	.80889E+04
9	.57646E+04	.46028E+04	.72196E+04
10	.57940E+04	.46063E+04	.72880E+04

Questions:

(a) (**20 points.**) Write an explanation of the above analyses as if you were reporting your conclusions to a sales manager. Your explanation should cover (but need not be limited to) (i) reasons for taking logarithms, (ii) what you learn from the ACF/PACF plots (iii) which analysis fits the data best, (iv) your recommendation of the future forecasts, (v) anything else which you think is important.

(b) (**5 points.**) The actual values of the series during January–October 1975 were:

2873 5556 5389 6135 6707 5220 6249 5281 4192 4867

In the light of this information, would you change any of your responses in (a)?

Go to next page for question 3.

3. (**25 points total.**) Figure 3 (page 10) shows annual mean temperatures in the northern hemisphere (NH), and the southern hemisphere (SH), as well as the Southern Oscillation Index (SOI). The SOI is closely connected with the meteorological phenomenon known as El Niño, which is well known to have a strong effect on global weather patterns. For ease of comparisons of covariances and regression coefficients, all three series have been scaled to have sample mean 0 and sample variance 1.

A 3-variable autoregressive model was fitted by the Yule-Walker method choosing the order of model by AICC. The results were:

```
Optimal value of p = 1

 PHI(0)
      -.000022
      -.000007
       .000061

 PHI(1)
       .556619       .358413       .120429
       .037792       .888829       .154141
       .230355      -.351434       .099649

Y-W White Noise Covariance Matrix, V
       .261731       .154665      -.186928
       .154665       .217977      -.186564
      -.186928      -.186564       .939917

AICC = .672657E+03
```

Forecasts up to 10 steps ahead have been computed from this model and are also plotted on Figure 3 (point forecasts shown by `x`, with 95% prediction bounds).

(a) (**10 points.**) What does the fitted model tell us about relationship among the three variables?

[A few hints here. Most of the earth's land mass is in the northern hemisphere, and temperatures change more slowly in the sea than on land. Also, although El Niño does affect global temperatures, SOI is primarily an index of atmospheric circulation over the Southern Hemisphere ocean. So we would expect some asymmetry between the NH and SH temperatures. What I am asking you is to say what you can deduce based on the above autoregressive model.]

(b) (**10 points.**) Suppose you wanted to calculate mean squared prediction error (MSPE) for forecasts of up to 5 steps ahead. *Without doing the actual calculations*, explain how you would calculate these MSPEs from the ITSM output.

(c) (**5 points.**) Do the forecasts look reasonable? If you had an opportunity to do the analysis a different way, what changes would you recommend?

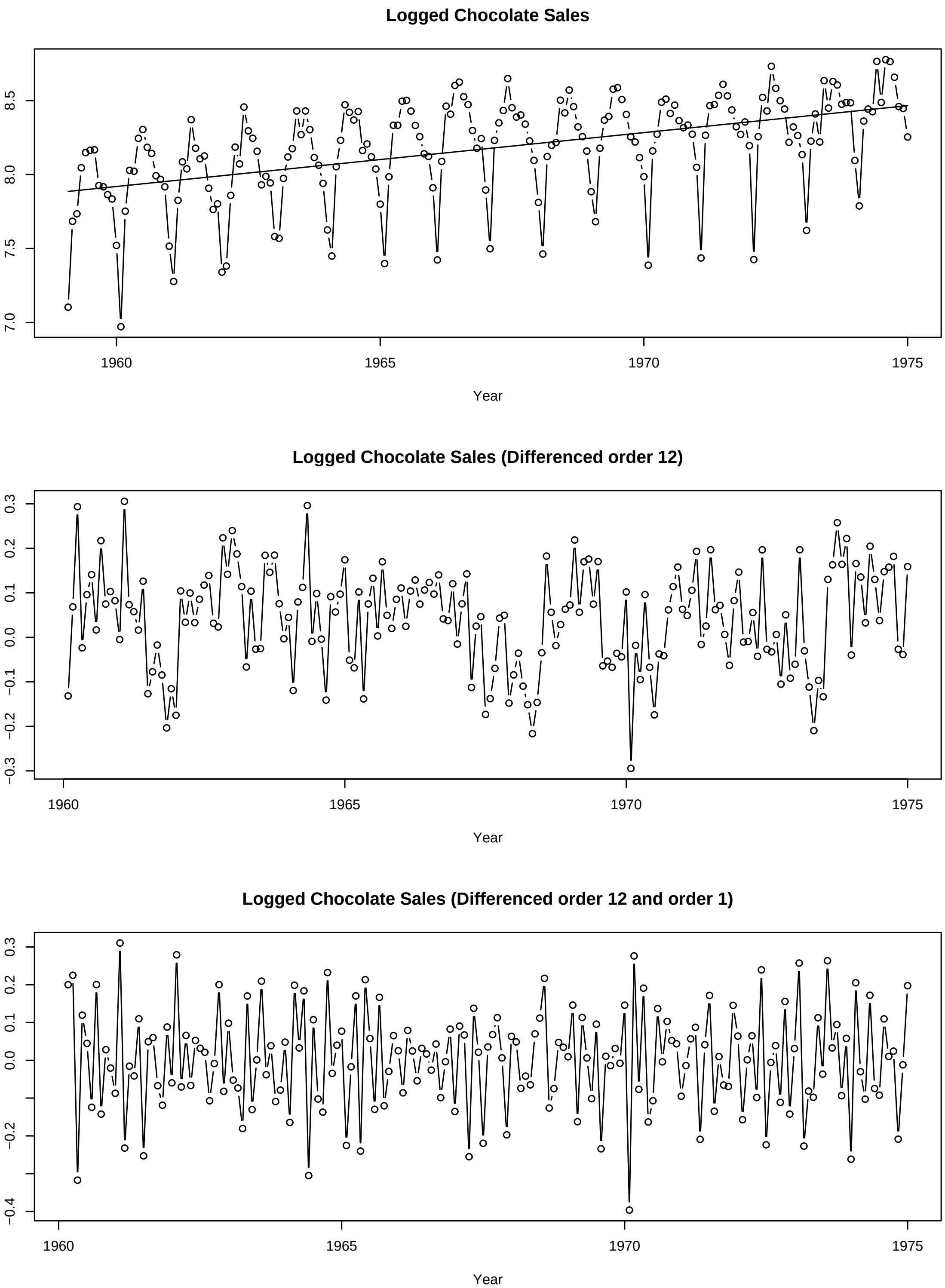

Figure 1: Logarithms of chocolate sales and differenced series. Linear trend shown on top plot.

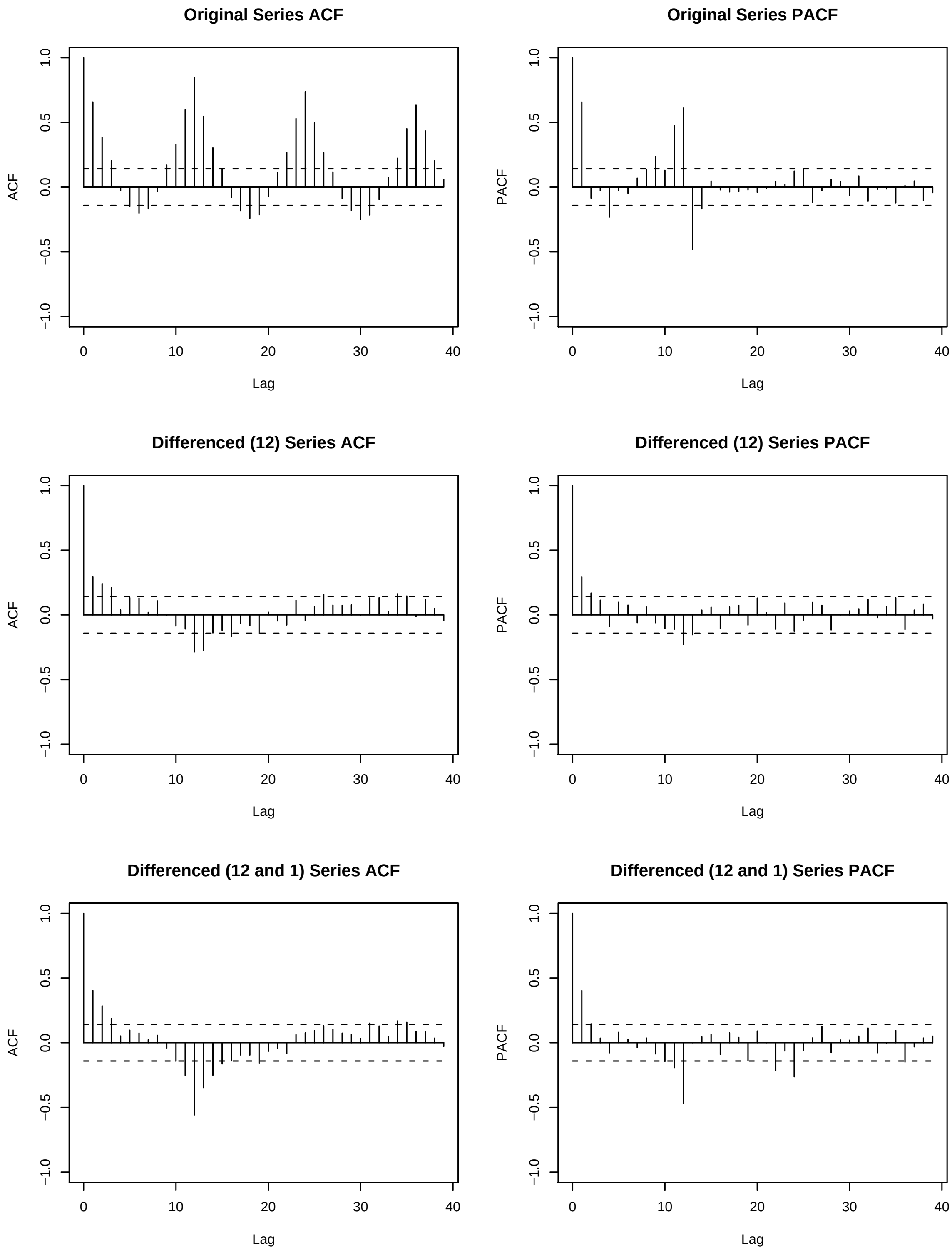

Figure 2: ACF/PACF of series in Figure 1.

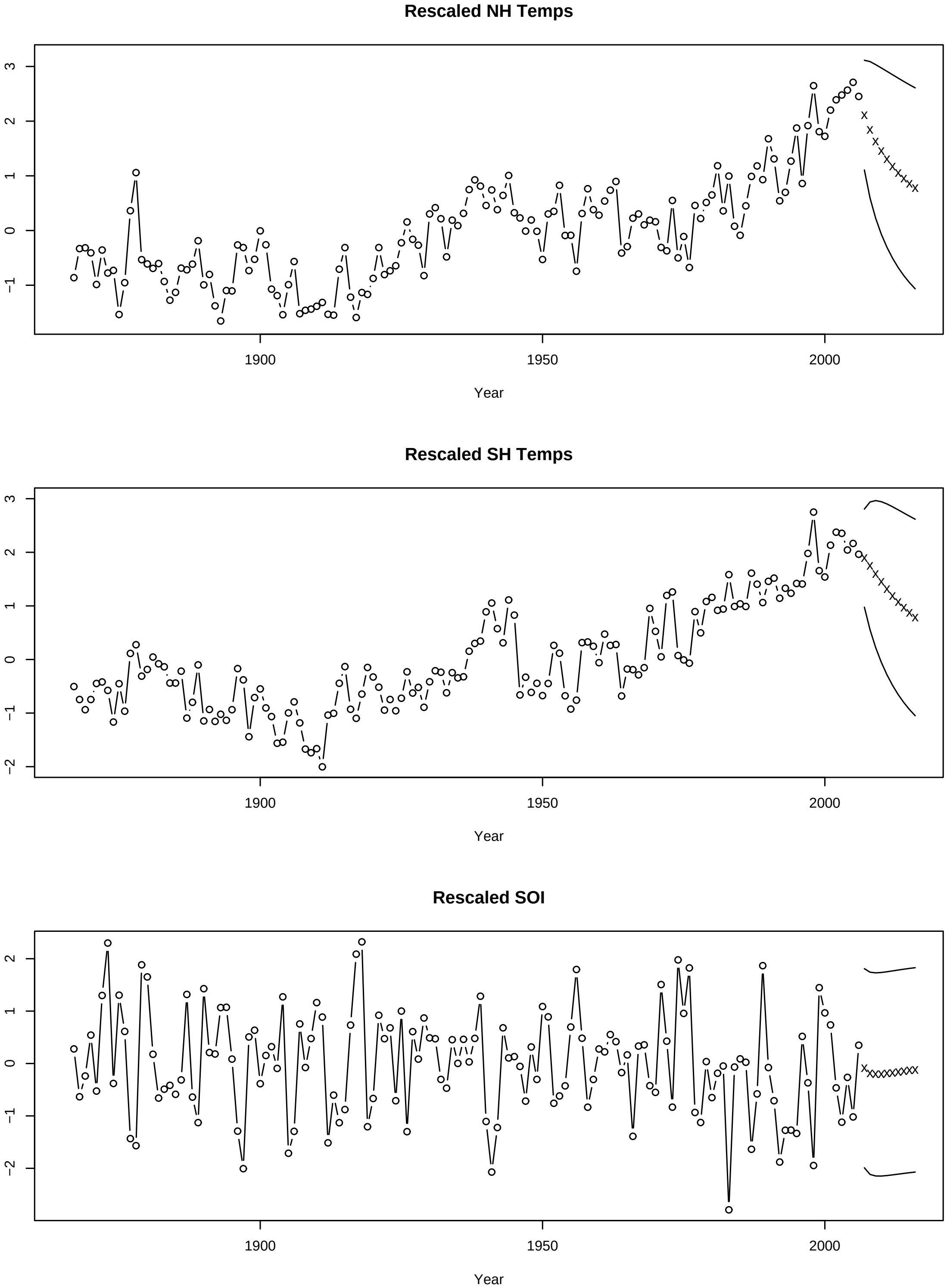

Figure 3: Top: Northern hemisphere mean temperatures, 1867–2006. Middle: Southern hemisphere mean temperatures, 1867–2006. Bottom: Southern Oscillation Index, 1867–2006. All three series have been rescaled to have mean 0 and variance 1. Also shown are forecasts up to 10 steps, and 95% prediction bounds.

SOLUTIONS

1. (a) $\phi(z) = 1 - 0.2z + 0.48z^2 = 0$ when $z = \frac{0.2 \pm \sqrt{0.2^2 - 4\times 0.48}}{2\times 0.48} = 0.208 \pm 1.428i$ for which $|z| = 1.443 > 1$. Also $\theta(z) = 1 - 0.7z$ with root at $z = \frac{1}{0.7}$ also satisfies $|z| > 1$. So the answer is yes to (i), (ii) and (iii).

   (b) $\psi(z) = \frac{1-0.7z}{1-0.2z+0.48z^2} = (1-0.7z)\{1+(0.2z-0.48z^2)+(0.2z-0.48z^2)^2+(0.2z-0.48z^2)^3+ ...\} = (1-0.7z)(1+0.2z-0.48z^2+0.04z-0.192z^3+0.008z^3+...) = (1-0.7z)(1+0.2z-0.44z^2-0.184z^3+...) = (1-0.5z-0.58z^2+0.124z^3+...)$. Note that we have used the expansion $\frac{1}{1-x} = 1+x+x^2+x^3+...$ and we have ignored all terms higher order than z^3. Thus $\psi_1 = -0.5,\ \psi_2 = -0.58,\ \psi_3 = 0.124$.

   (c) $\pi(z) = \frac{1-0.2z+0.48z^2}{1-0.7z} = (1-0.2z+0.48z^2)(1+0.7z+0.49z^2+0.343z^3+0.2401z^4...) = (1+0.5z+0.83z^2+0.581z^3+0.4067z^4+...)$. So $\pi_1 = 0.5,\ \pi_2 = 0.83,\ \pi_3 = 0.581, \pi_4 = 0.4067$.

   [*Comment about (b) and (c)*: The above is the "from first principles" solution. It is also possible to use formulas (3.1.7), (3.1.8), pp. 85–86, of the course text.]

   (d) $\tilde{P}_t X_{t+1} = -\sum_{j=1}^{\infty} \pi_j X_{t+1-j} = -0.5X_t - 0.83X_{t-1} - 0.581X_{t-2} - ....$

   $\tilde{P}_t X_{t+2} = -\pi_1 \tilde{P}_t X_{t+1} - \sum_{j=2}^{\infty} \pi_j X_{t+2-j} = -0.5(-0.5X_t - 0.83X_{t-1} - 0.581X_{t-2} - ...) - 0.83X_t - 0.581X_{t-1} - 0.4067X_{t-2}... = -0.58X_t - 0.166X_{t-1} - 0.1162X_{t-2} - ...$

   (e) The general formula for the MSPE of $\tilde{P}_t X_{t+h}$ is $\sigma^2 \sum_{j=0}^{h-1} \psi_j^2$ so the specific answers for $h = 1, 2, 3$ are (i) $\sigma^2 = 3$, (ii) $\sigma^2(1+\psi_1^2) = 3(1+0.5^2) = 3.75$, (iii) $\sigma^2(1+\psi_1^2+\psi_2^2) = 3(1+0.5^2+0.58^2) = 4.7592$.

   [*Comment about (d) and (e):* Some students erroneously used the prediction formulas for AR(2), which lead to $\tilde{P}_t X_{t+1} = 0.2X_t - 0.48X_{t-1}$, $\tilde{P}_t X_{t+2} = -0.44X_t - 0.096X_{t-1}$. However, some students who did this also calculated correctly the MSPEs of these predictors, which are 4.47, 3.81, 5.04 for $h = 1, 2, 3$ (the $h = 3$ case being also based on the AR(2) predictor). So while this wasn't the correct answer, I was willing to give significant partial credit to students who followed this method consistently.]

   (f) Taking covariances of both sides of the model equation with X_{t-k} for $k \geq 0$,

$$
\begin{aligned}
\gamma_X(k) - 0.2\gamma_X(k-1) + 0.48\gamma_X(k-2) &= \mathrm{Cov}(Z_t - 0.7Z_{t-1}, \sum_{j=0}^{\infty} \psi_j Z_{t-k-j} \\
&= \begin{cases} (1-0.7\psi_1)\sigma^2 & \text{if } k = 0, \\ -0.7\sigma^2 & \text{if } k = 1, \\ 0 & \text{if } k \geq 2. \end{cases}
\end{aligned}
$$

   Hence

$$
\begin{aligned}
\gamma_X(0) - 0.2\gamma_X(1) + 0.48\gamma_X(2) &= 4.05, \\
\gamma_X(1) - 0.2\gamma_X(0) + 0.48\gamma_X(1) &= -2.1, \\
\gamma_X(2) - 0.2\gamma_X(1) + 0.48\gamma_X(0) &= 0, \\
\gamma_X(3) - 0.2\gamma_X(2) + 0.48\gamma_X(1) &= 0.
\end{aligned}
$$

   These are four simultaneous equations in four unknowns; we solve these equations and then define $\rho_X(h) = \gamma_X(h)/\gamma_X(0)$ for $h = 1, 2, 3$.

   [The numerical solutions are $\gamma_X(0) = 5.1650498$, $\gamma_X(1) = -0.7209392$, $\gamma_X(2) = -2.6234117$, $\gamma_X(3) = -0.1786315$ which lead to $\rho_X(1) = -0.13958030$, $\rho_X(2) = -0.50791606$, $\rho_X(3) = -0.03458467$.]

2. (a) There is no "perfect solution" to this but following are an outline of what I would see as the main points — credit given will depend on how well you make your own arguments to support these or other points.

For reference I'll use the shorthand S1, S2, S3 to refer to the original series (S1), the series differenced at lag 12 (S2) and the series differenced at lags 12 and 1 (S3).

(i) Most likely, the reason for taking logarithms was to stabilize the amplitude of the seasonal cycle.

(ii) The ACF for S1 shows strong seasonality and very slow decay to 0 — indicates nonstationarity. For S2, ACF and PACF both have significant values around lag 12 or 13, but not for higher lags. For S3, the ACF shows larger (in absolute value) values over lags 1–15 and both ACF and PACF show some significant values at much higher lags. On the basis of that, S2 seems the best bet to model as a stationary series.

(iii) For S1, the fitted model ARMA(12,12) is very high order (which must raise some doubts over whether it has been correctly identified) and the Ljung-Box statistic rejects the selected model (which strengthens that doubt; however, the other residual tests are OK). For S3, Ljung-Box is very highly significant, which confirms that the ACFs at high lags (> 13) are really a problem with fitting this model. (The turning points test also rejects though it is less clear how to interpret this.) It also looks as though the MA coefficients at lags 1, 2, 4, 7, 8, 9, 12 and 13 are all significant, and while this does not in itself create reason to doubt the model, it does mean that the MA(13) model is cumbersome and difficult to interpret. For S2, it is again true that several of the MA(13) coefficients are significant, but the model passes all the residual tests, and this is also the best of the three models when assessed by either AICC or BIC. Conclusion: although none of the three models is perfect, S2 seems the best fitting.

(iv) I would use the forecasts fitted to S2.

(b) For S1, the future values that lie outside the prediction limits are steps 2, 4 and 9. For S2, steps 1, 2, 4, 9. For S3, 1 (just), 2, 8, 9. While this would not make me change my recommendation, it does show that the three sets of forecasts perform about equally well in terms of actual forecast accuracy.

3. (a) Again, this question is rather open-ended so credit will depend on how well you manage to make pertinent points rather than anything specific. Some things you might pick up are: (i) all the entries of ϕ_0 are very close to 0 — but this is expected, because we subtracted the means; (ii) the covariance matrix Σ (V in the ITSM notation) shows apparently significant correlations among all three variables, but the correlations between SOI and either NH or SH are negative — so presumably, temperatures are high when SOI is low. (iii) As for the Φ_1 matrix, one thing you might point out is that the coefficient of $X_{t,2}$ on $X_{t-1,1}$ (.037792) is low — in other words, last year's NH temprature does not have much influence on SH temperature — this may be a consequence of the slow change of temperature over the ocean, as suggested by the hint. Athough no formal significance tests are provided, from the magnitude of the coefficients it looks as though all the other regression coefficients are $\neq 0$, which is a little surprising but again confirms that all three variables seem interrelated. [The other thing that is surprising is that the coefficients of $X_{t,3}$ on either $X_{t-1,1}$ and $X_{t-1,2}$ seem to be significant — in other words, last year's temperatures seem to affect this year's El Niño. This may be an artifact of the way the data were calculated.]

(b) The relevant formula for the prediction error covariance of the h-step prediction is

$\sum_{j=0}^{h-1} \Psi_j \Sigma \Psi_j^T$ (course notes part IV, page 8) where for the AR(1) process, $\Psi_j = \Phi_1^j$ (course notes part IV, page 6). So for example, for $h = 5$ the covariance formula is

$$\Sigma + \Phi_1 \Sigma \Phi_1^T + \Phi_1^2 \Sigma (\Phi_1^2)^T + \Phi_1^3 \Sigma (\Phi_1^3)^T + \Phi_1^4 \Sigma (\Phi_1^4)^T$$

where $\Phi_1 = \begin{pmatrix} .556619 & .358413 & .120429 \\ .037792 & .888829 & .154141 \\ .230355 & -.351434 & .099649 \end{pmatrix}$, $\Sigma\ (= V) = \begin{pmatrix} .261731 & .154665 & -.186928 \\ .154665 & .217977 & -.186564 \\ -.186928 & -.186564 & .939917 \end{pmatrix}$ from the ITSM output. Thus, we multiply and add these matrices; the individual MSPEs are the diagonal entries of the covariance matrices.

(c) The one thing missing from our model is that we haven't made any attempt to represent the trend (including SOI as a covariate does not help with the trend, because SOI itself has no trend). This is seen in the forecasts, which seem to be trying to bring the temperature series back to their mean. This would be correct if there was indeed no trend, but given the strong visual evidence of a trend (supplemented by all the scientific knowledge we have about the effects of greenhouse gases, etc.) the logical conclusion is that the model is wrong, and therefore the forecasts cannot be trusted. [Extra credit for pointing out that maybe this is why some of the answers in part (a) didn't seem to make sense either.] Possible remedies: (i) Fit an alternative model that accounts explicitly for the trend in NH and SH temperatures (e.g. regression or differencing), (ii) introduce some other covariate that might explain the trend, such as an indicator of carbon dioxide concentrations in the atmosphere.

Comments on the scores. For the final, mean score was 80.4; SD 11.8; 5-number summary (54, 76, 83, 90, 98). Question 1 was, surprisingly, the best answered of the three, with a median score of 44/50, as against 21/25 for question 2 and 19/25 for question 3. I have the feeling that maybe students were running out of time and therefore didn't look at Q3 in as much detail as the other two questions. In Q1, several students treated it as an AR(2) process, ignoring the fact that it was actually ARMA(2,1) — this affected some of the answers to parts (d), (e) and (f). Q2 was also generally thoroughly answered, though for part (b), some students lost points for not being specific about their answers (e.g. if the reason for either changing or not changing your mind was the number of true observations that lay in the respective prediction intervals, you should be explicit about which observations you are talking about). In question 3(a), a complete answer needed to discuss both the Φ_1 and the V (or Σ) matrix — not everyone did that. Overall, however, I felt that this was not an easy exam and students did well at it — that was my reason for giving a higher percentage of A and B grades than I normally do in undergraduate classes.