CHAPTER 4

STATISTICAL MODELLING

In this chapter, the results of the model fitting are presented. These results may be classified as follows.

(a) Analysis of the times of occurrence of successive high and low tides at Pattani and Songkla during 1996, based on the tide tables.

(b) Analysis of the heights of successive high and low tides at Pattani and Songkla during 1996, based on the tide tables.

In the preliminary analysis we saw (Figures 14 and 15) that the times of occurrence of each high and low tide, relative to the lunar day, separate into four clusters approximately 6 hours and 12 minutes apart. This is a standard feature of semidiurnal tide patterns.

We also saw in the preliminary analysis that the variation in water levels is complicated, showing seasonal variation throughout the year (with peaks in January and December and troughs in the July-August period (Figures 12 and 13), as well as monthly and daily cycles (Figures 10 and 11).

The daily cycle in the water levels is removed if we separate both the high and low tides into those occurring at approximately the same time each lunar day. This means that instead of analysing two series of water levels at each location (one for the high tide and one for the low tide), we analyse four series as follows.

(i) the first low tide, which occurs at the beginning of the lunar day;

(ii) the first high tide, which occurs approximately 6 hours and 12 minutes into the lunar day;

(iii) the second low tide, occurring at approximately 12 hours and 25 minutes;

(iv) the second high tide, occurring at approximately 18 hours and 37 minutes.
In this chapter, the methods of time series analysis, using harmonic periodic components to describe the signal, and autoregressive parameters to describe the

residuals, are applied to the eight series corresponding to the four tides, as described above, at the two locations.

1. Times of occurrence at Pattani

Figure 16 shows the result of fitting the model to the series of times of occurrence of the first low tide at Pattani during 1996.

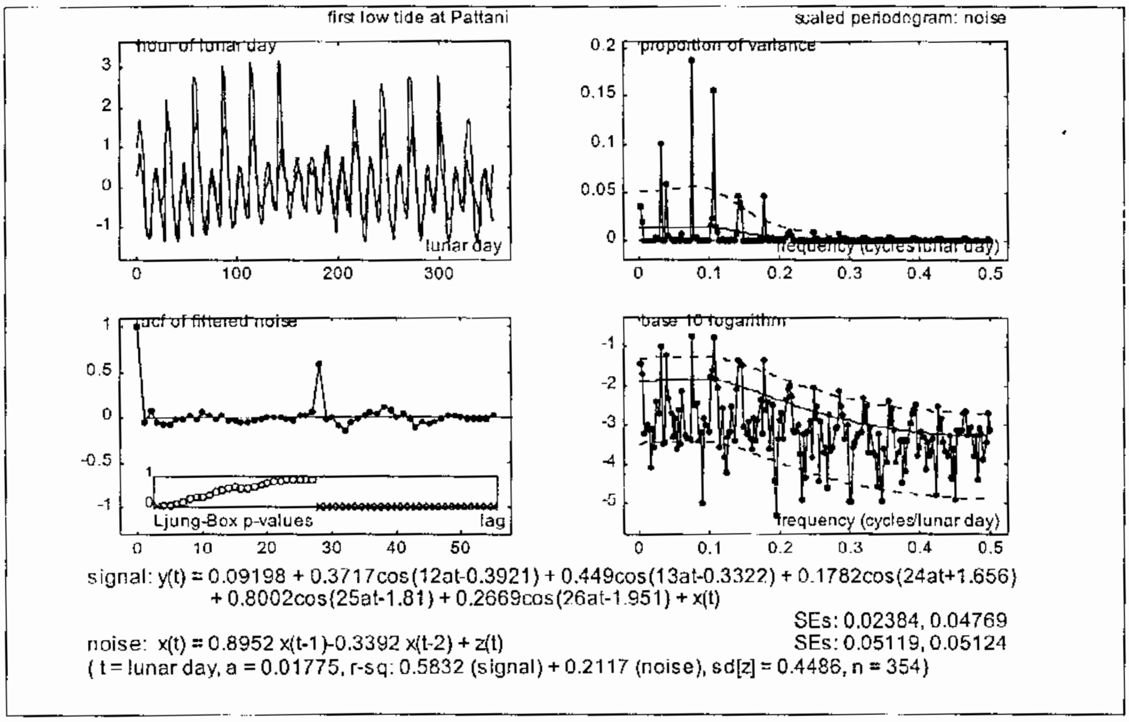


Figure 16: Analysis of times of occurrence of the first low tide at Pattani in 1996

The fitted signal includes harmonics at 12, 13, 24, 25 and 26 cycles with amplitudes 0.37, 0.45, 0.18, 0.80 and 0.27, respectively. These correspond to the movement of the moon around the earth. The fit is reasonable but not particularly good. The r-squared for the signal is only 0.5603. Judging from the graph of the time series with the model superimposed, the model fails to accommodate the high single peaks in the data which occur once during each lunar revolution.

After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 0.95 and -0.32. The r-squared associated with this model is 0.233, giving a total r-squared of 0.793.

The series of residuals, after removing the signal and filtering the noise based on the autoregressive process, still contains a statistically significant autocorrelation at lag 28.

Figure 17 shows the result of fitting the model to the series of times of occurrence of the first high tide at Pattani.

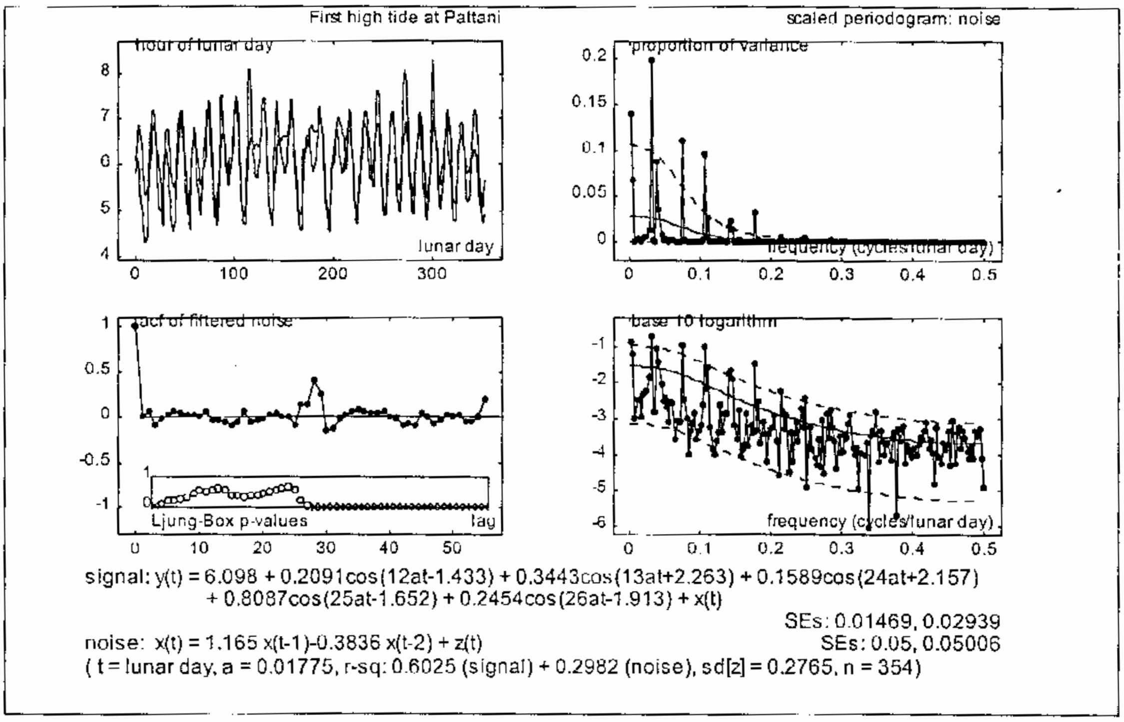


Figure 17: Analysis of times of occurrence of the first high tide at Pattani in 1996

The fitted signal includes harmonics at 12, 13, 24, 25 and 26 cycles with amplitudes 0.21, 0.34, 0.16, 0.81 and 0.25, respectively. These again correspond to the movement of the moon around the earth, and as for the first low tide the fit is reasonable but not particularly good. The r-squared for the signal is 0.6025. Judging from the graph of the time series with the model superimposed, the model fails to accommodate some of the peaks and troughs in the data which occur during successive lunar revolutions.

After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.165 and -0.38. The r-squared associated with this model is 0.298, giving a total r-squared of 0.900.

As in the case of the first low tide occurrence times, the filtered noise has a significant autocorrelation at lag 28.

Figure 18 shows the analysis for the times of occurrence of the second low tide at Pattani.

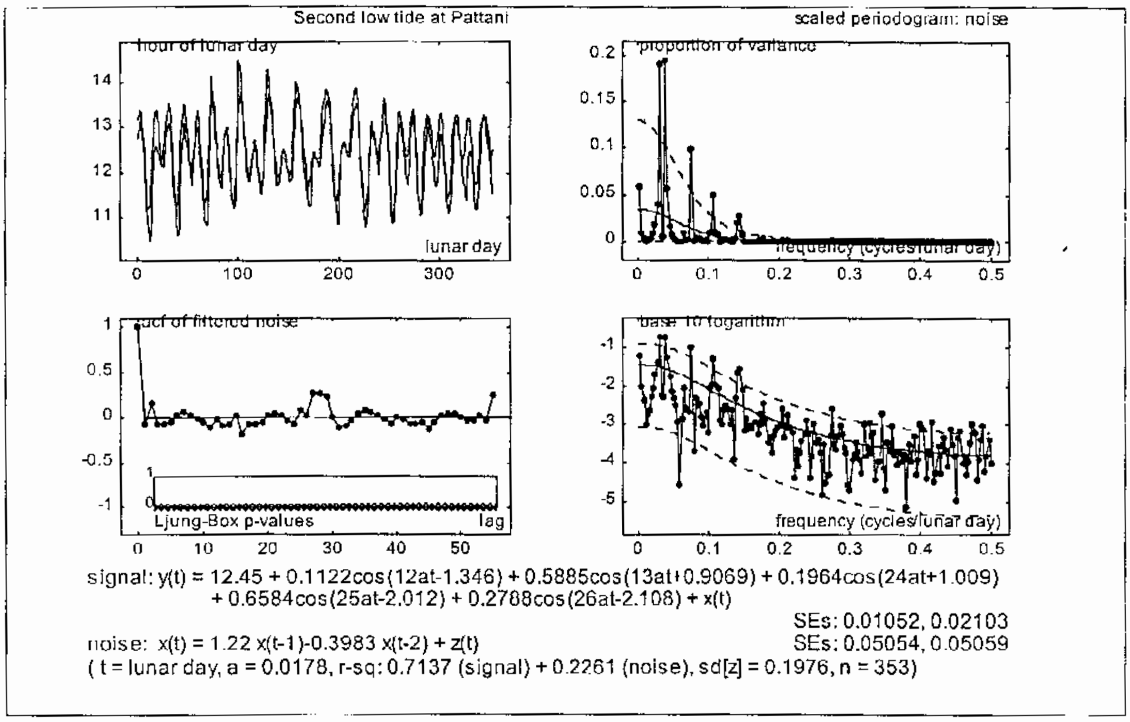


Figure 18: Analysis of times of occurrence of the second low tide at Pattani in 1996

The fitted signal includes harmonics at 12, 13, 24, 25 and 26 cycles with amplitudes 0.11, 0.59, 0.20, 0.66 and 0.28, respectively. As before these correspond to the movement of the moon around the earth. Again, the fit is reasonable but not particularly good. The r-squared for the signal is 0.7137.

After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.22 and -0.40. The r-squared associated with this model is 0.226, giving a total r-squared of 0.94. The Ljung-Box p-values are all smaller than 0.05, indicating that there is statistically significant autocorrelation in the filtered series.

Figure 19 shows the analysis for the times of occurrence of the second high tide at Pattani.

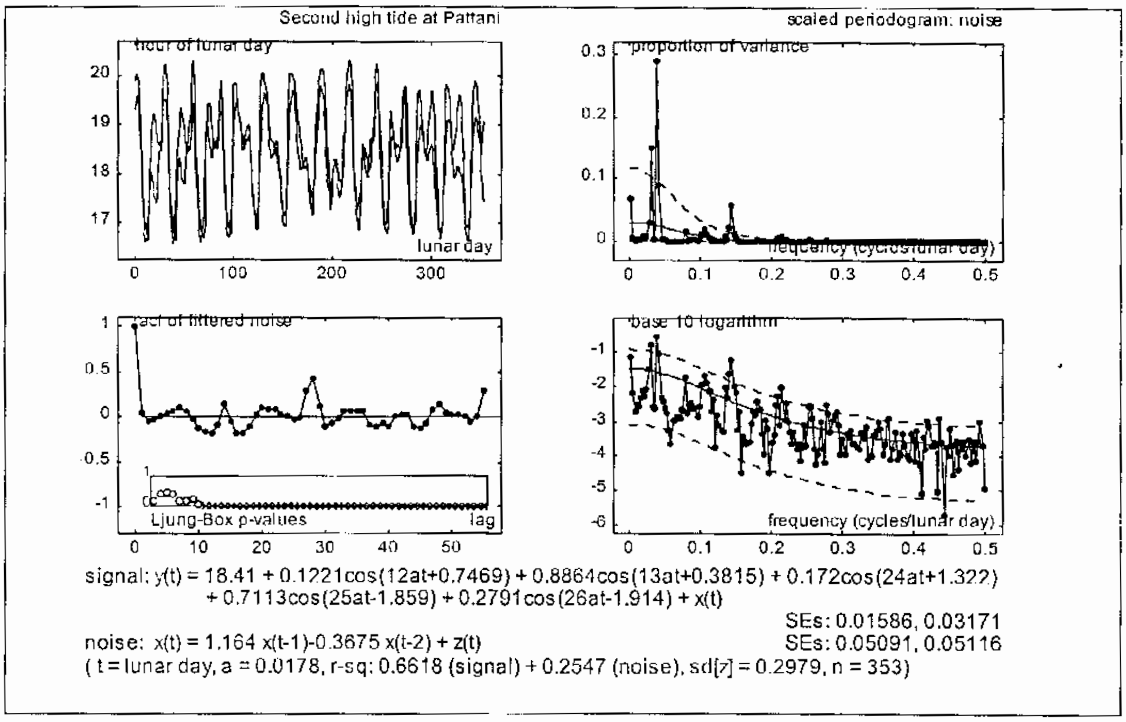


Figure 19: Analysis of times of occurrence of the second high tide at Pattani in 1996

The fitted signal includes harmonics at 12, 13, 24, 25 and 26 cycles with amplitudes 0.12, 0.89, 0.17, 0.71 and 0.28, respectively. These again correspond to the movement of the moon around the earth. As before, the fit is reasonable but not particularly good. The r-squared for the signal is 0.6618.

After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.164 and -0.37. The r-squared associated with this model is 0.2547, giving a total r-squared of 0.9165. The autocorrelation function has significant components at lags greater than 10, indicating the inadequacy of the model.

2. Times of occurrence at Songkhla

Figure 20 shows the analysis for the times of occurrence of the first low tide at Songkhla.

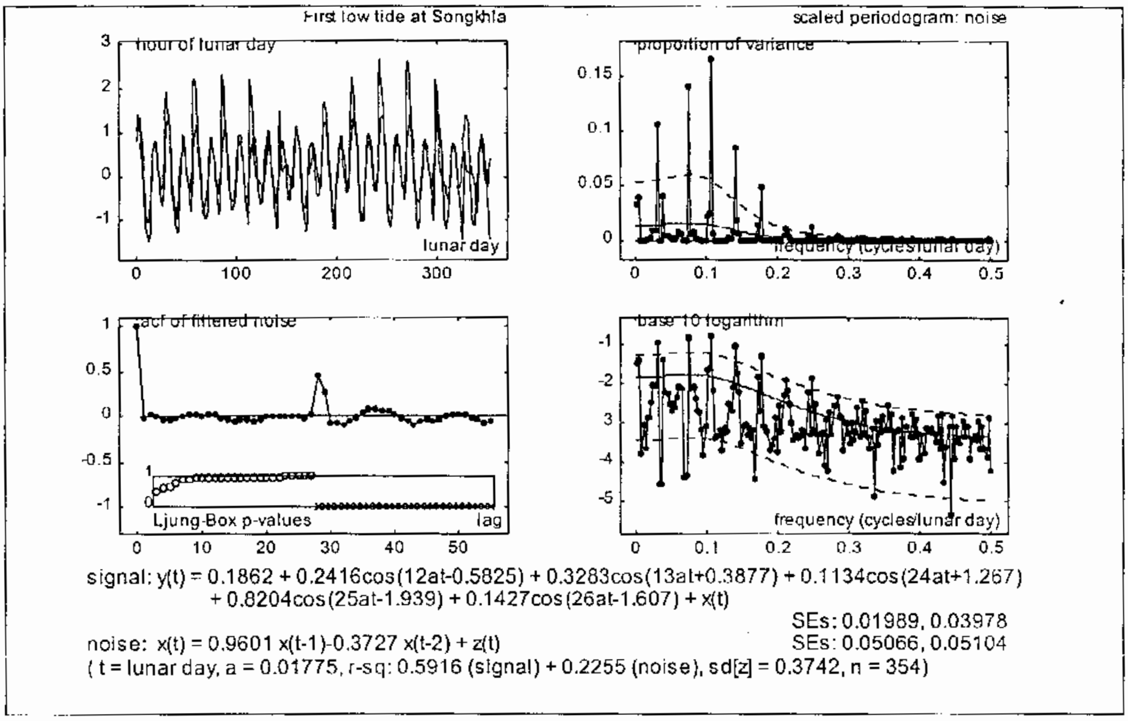


Figure 20: Analysis of times of occurrence of the first low tide at Songkhla in 1996

The fitted signal includes harmonics at 12, 13, 24, 25 and 26 cycles with amplitudes 0.24, 0.33, 0.11, 0.82 and 0.14, respectively. These correspond to the movement of the moon around the earth. The fit is reasonable but not particularly good. The r-squared for the signal is only 0.5916. Judging from the graph of the time series with the model superimposed, the model fails to accommodate the peaks in the data, which occur during each lunar revolution.

After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 0.96 and -0.37. The r-squared associated with this model is 0.2255, giving a total r-squared of 0.8171.

The series of residuals, after removing the signal and filtering the noise base on the autoregressive process, still contains a statistically significant autocorrelation at lag 28.

Figure 21 shows the analysis for the times of occurrence of the first high tide at Songkhla.

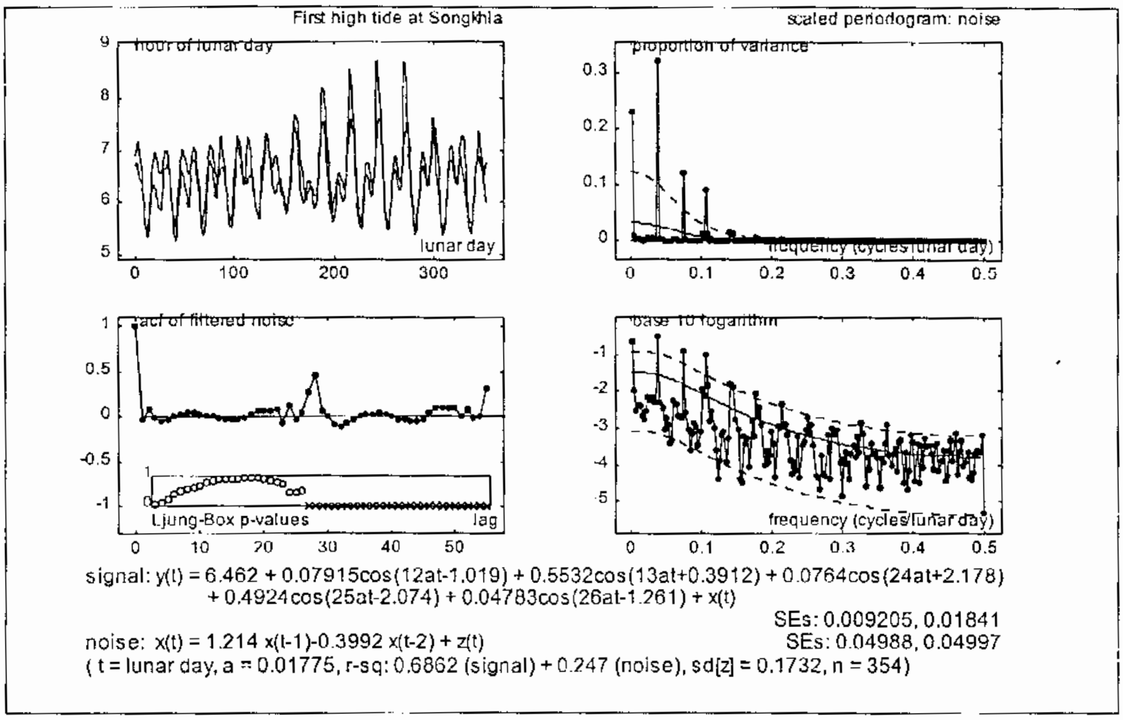


Figure 21: Analysis of times of occurrence of the first high tide at Songkhla in 1996

The fitted signal includes harmonics at 12, 13, 24, 25 and 26 cycles with amplitudes 0.08, 0.55, 0.08, 0.50 and 0.05, respectively. These correspond to the movement of the moon around the earth. Again, the fit is reasonable but not particularly good. The r-squared for the signal is 0.6862. Judging from the graph of the time series with the model superimposed, the model fails to accommodate the peaks in the data, which occur during each lunar revolution.

After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.21 and -0.40. The r-squared associated with this model is 0.247, giving a total r-squared of 0.9332.

As in the case of the first low tide occurrence times, the filtered noise has a significant autocorrelation at lag 28.

Figure 22 shows the analysis for the times of occurrence of the second low tide at Songkhla.

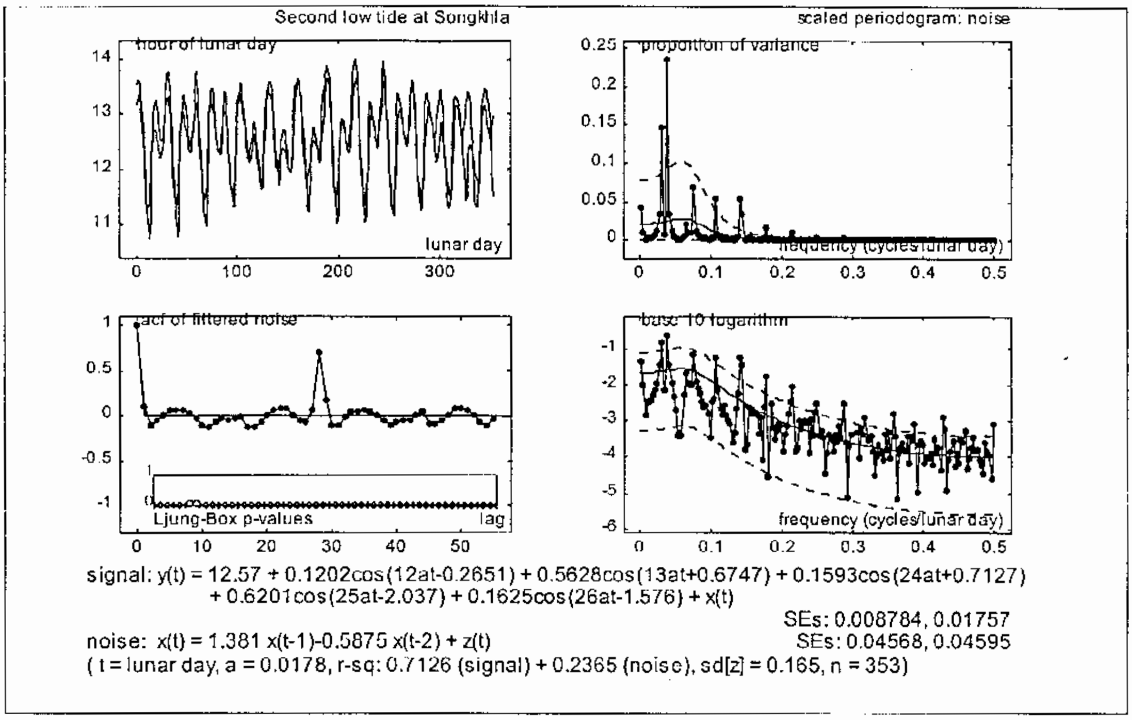


Figure 22: Analysis of times of occurrence of the second low tide at Songkhla in 1996

The fitted signal includes harmonics at 12, 13, 24, 25 and 26 cycles with amplitudes 0.12, 0.56, 0.16, 0.62 and 0.16, respectively. As before these correspond to the movement of the moon around the earth. Again, the fit is reasonable but not particularly good. The r-squared for the signal is 0.7126.

After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.38 and -0.59. The r-squared associated with this model is 0.2365, giving a total r-squared of 0.9491.

The Ljung-Box p-values are greater than 0.05 only for lag 7 and 8 indicating that the stochastic part of the model is inadequate. Also there is a high autocorrelation at lag 28.

Figure 23 shows the analysis for the times of occurrence of the second high tide at Songkhla.

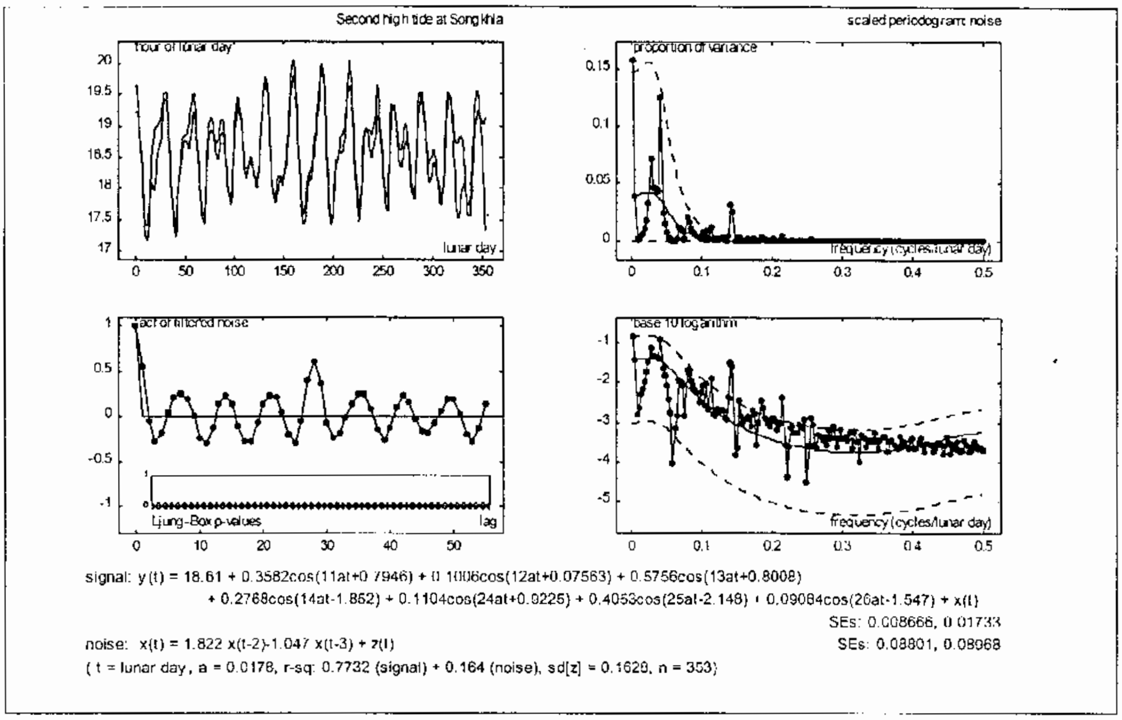


Figure 23: Analysis of times of occurrence of the second high tide at Songkhla in 1996

The fitted signal includes harmonics at 11, 12, 13, 14, 24, 25 and 26 cycles with amplitudes 0.358, 0.101, 0.576, 0.277, 0.110, 0.405 and 0.091, respectively. As before these correspond to the movement of the moon around the earth. The fit is not particularly good. The r-squared for the signal is 0.773.

After subtracting the signal, the noise may be modelled as a third order autoregressive process with parameters 0, 1.822 and -1.047. The r-squared associated with this model is 0.164, giving a total r-squared of 0.937.

The Ljung-Box p-values are all smaller than 0.05, indicating that there is statistically significant autocorrelation in the filtered noise series. The model is clearly not satisfactory, and a much more complicated model is required to obtain a reasonable fit.

3. Heights of high and low tides at Pattani

Figure 24 shows the analysis for the heights of the first high tide at Pattani.

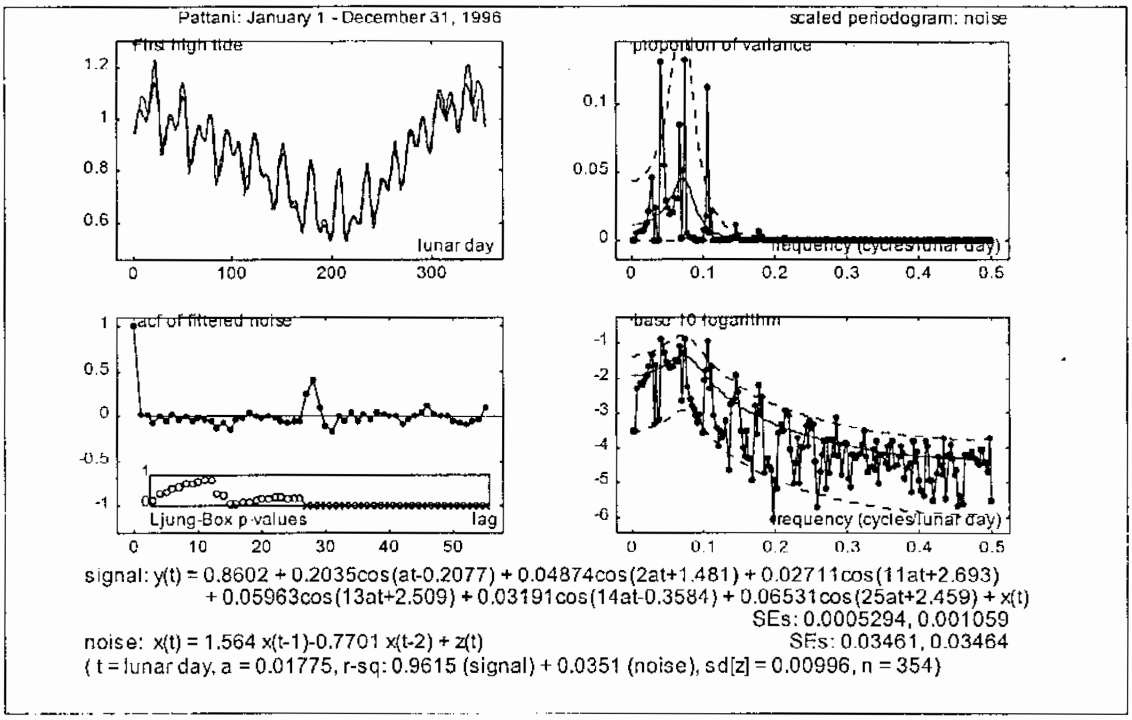


Figure 24: Analysis of heights of the first high tide at Pattani in 1996

The graph of the base 10 logarithm of the periodogram gives 95% confidence intervals for the individual periodogram values. Most of the periodogram values are inside these confidence limits which indicates that the model fits the data. The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.204, 0.049, 0.027, 0.060, 0.032 and 0.065, respectively. The first and second harmonics indicate a seasonal pattern, which may be caused by the monsoon winds. The other two harmonics correspond to the movement of the moon around the earth. The r-squared for the signal is 0.962. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.564 and -0.770. The r-squared associated with this model is 0.035, giving a total r-squared of 0.997, indicating that 99.7% of variation in the data can be explained by this model. The fit is thus extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values, indicated by circles on the graph, are greater than 0.05 only for lags less than 26. Also there is a high autocorrelation at lag 28. This shows that despite the high r-squared, some pattern remains in the residual filtered noise series.

Figure 25 shows the analysis for the heights of the first low tide at Pattani.

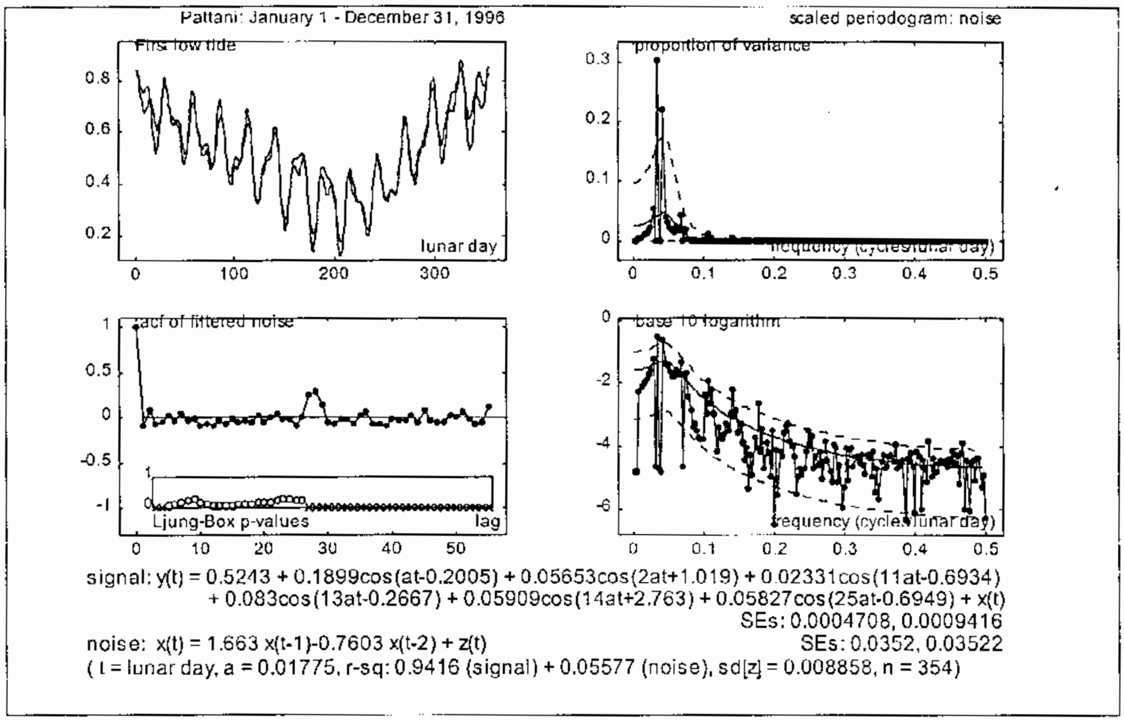


Figure 25: Analysis of heights of the first low tide at Pattani in 1996

The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.189, 0.056, 0.023, 0.083, 0.059 and 0.058, respectively. The r-squared for the signal is 0.942. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.65 and -0.76. The r-squared associated with this model is 0.056, giving a total r-squared of 0.998, indicating that 99.8% of variation in the data can be explained by this model. The fit is thus extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values, indicated by circles on the graph, are greater than 0.05 only for lags less than

26. Also there is a high autocorrelation at lag 28. This shows that despite the high r-squared, some pattern remains in the residual filtered noise series.

Figure 26 shows the analysis for the heights of the second high tide at Pattani.

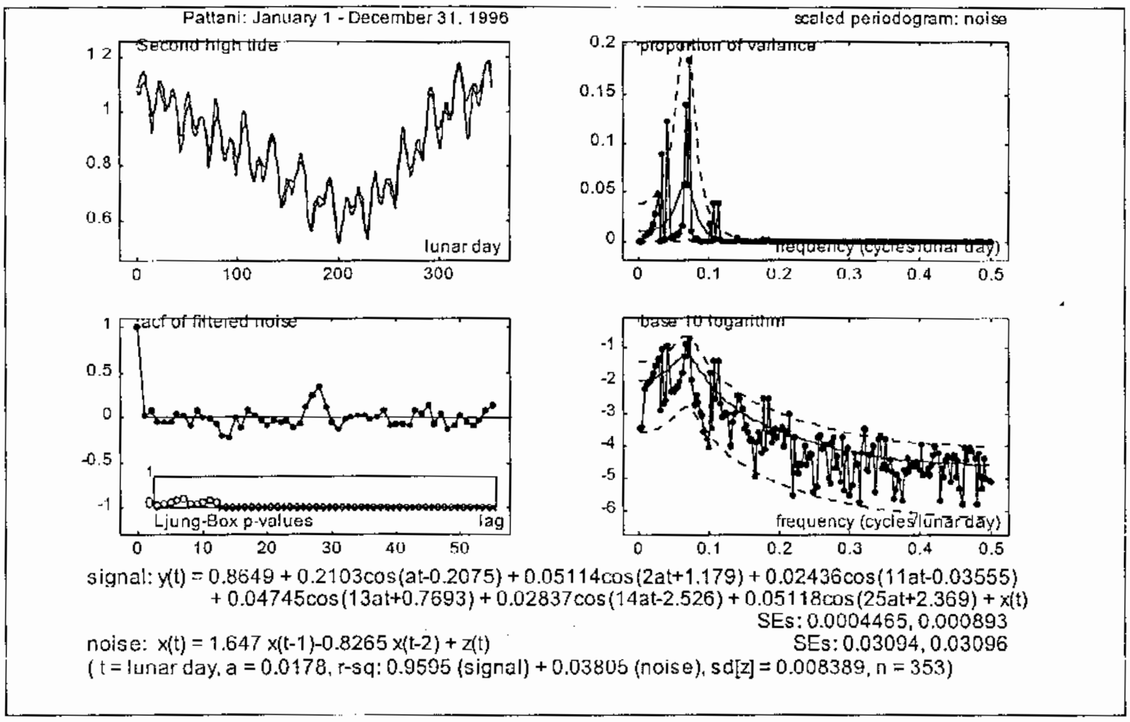


Figure 26: Analysis of heights of the second high tide at Pattani in 1996

The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.210, 0.051, 0.024, 0.047, 0.028 and 0.051, respectively. The r-squared for the signal is 0.960. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.65 and -0.83. The r-squared associated with this model is 0.038, giving a total r-squared of 0.998, indicating that 99.8% of variation in the data can be explained by this model. As before the fit is extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values, indicated by circles on the graph, are greater than 0.05 only for lags less than 12. Also there is a high autocorrelation at lag 28. This shows that despite the high r-squared, some pattern remains in the residual filtered noise series.

Figure 27 shows the analysis for the heights of the second low tide at Pattani.

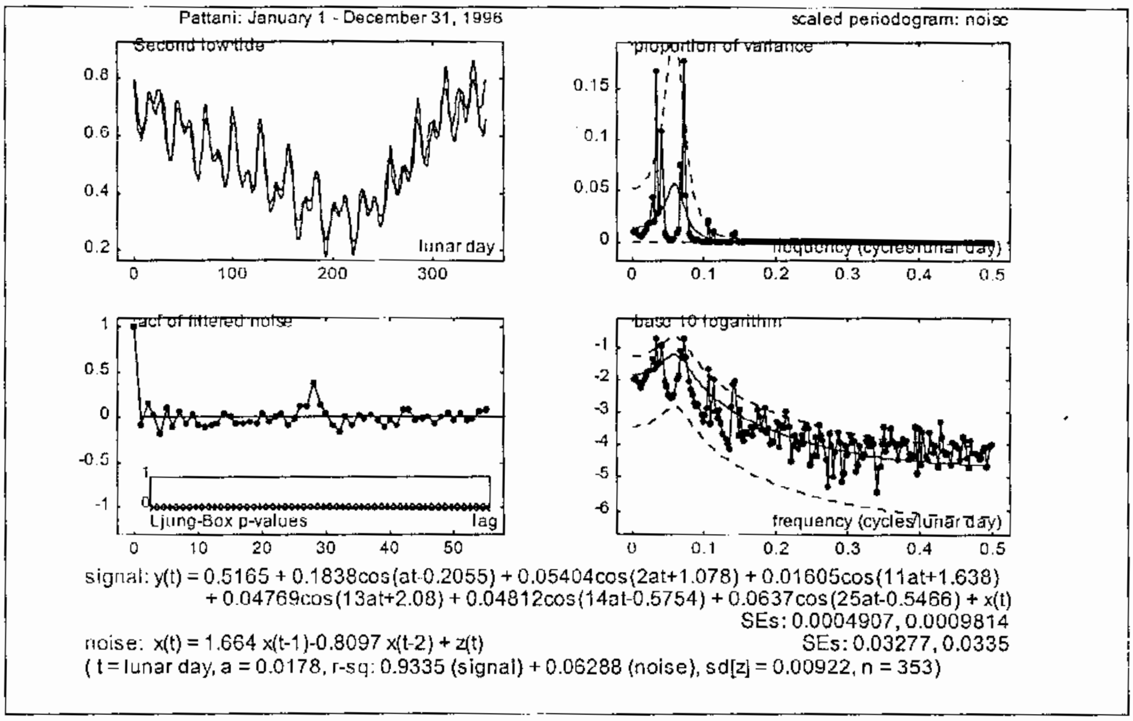


Figure 27: Analysis of heights of the second low tide at Pattani in 1996

The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.184, 0.054, 0.016, 0.048, 0.048 and 0.064, respectively. The r-squared for the signal is 0.934. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.66 and -0.81. The r-squared associated with this model is 0.063, giving a total r-squared of 0.997, indicating that 99.7% of variation in the data can be explained by this model. The fit is again extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values are all smaller than 0.05, indicating that there is substantial statistically significant autocorrelation in the filtered series.

4. Heights of high and low tides at Songkhla

Figure 28 shows the analysis for the heights of the first high tide at Songkhla.

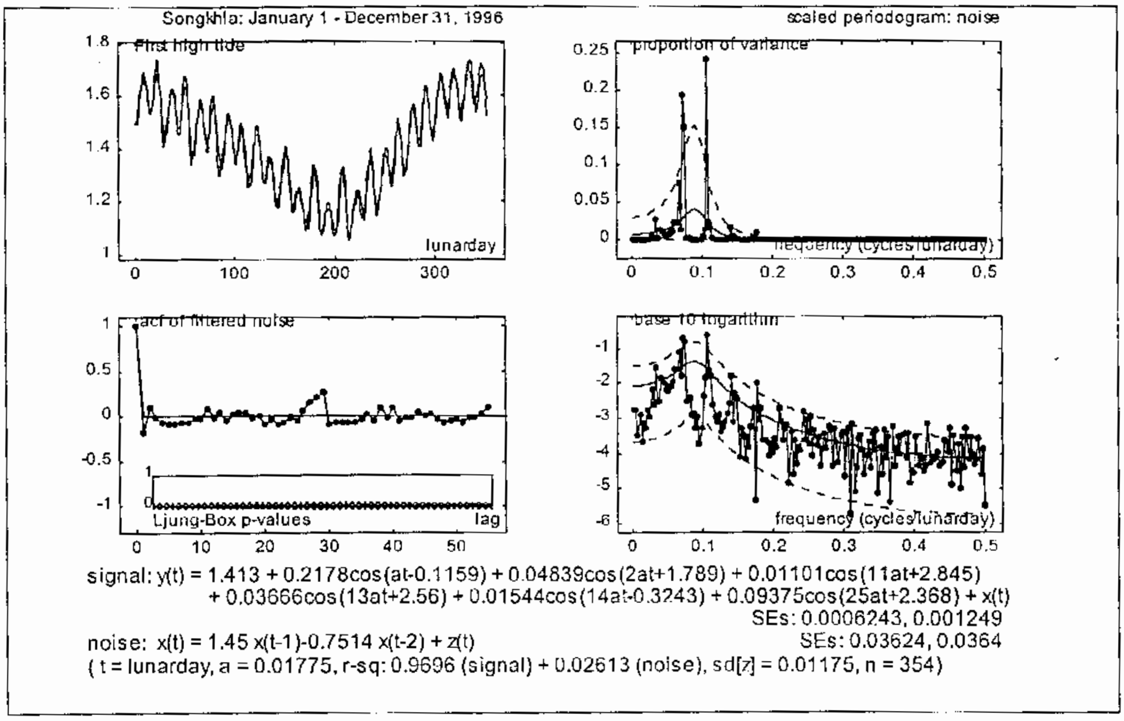
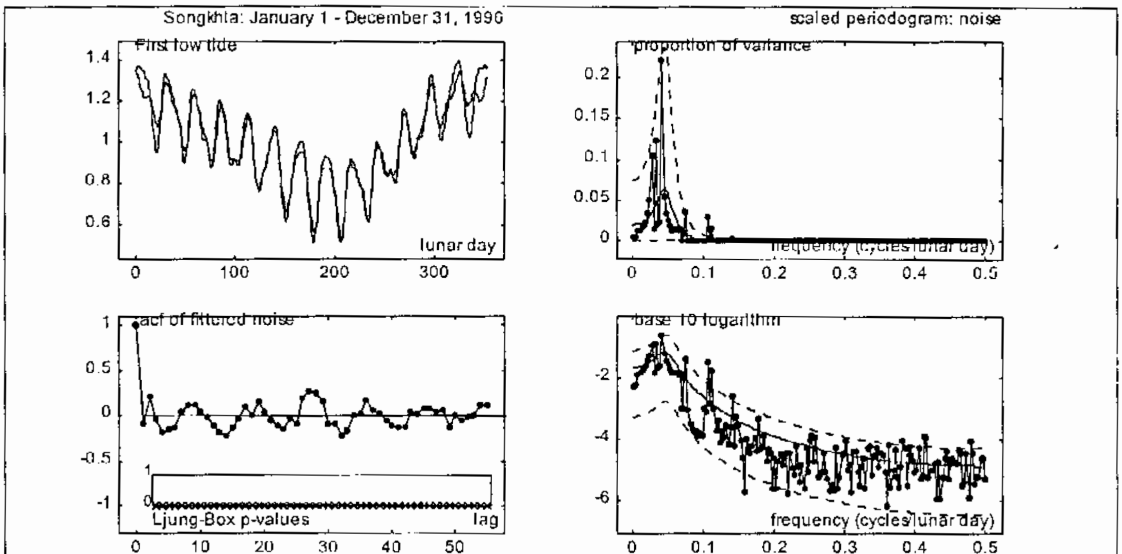


Figure 28: Analysis of heights of the first high tide at Songkhla in 1996

The graph of the base 10 logarithm of the periodogram gives 95% confidence intervals for the individual periodogram values. The results are similar to those for Pattani. Most of the periodogram values are inside these confidence limits which indicates that the model fits the data. The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.218, 0.048, 0.011, 0.037, 0.015 and 0.094, respectively. These correspond to the movement of the moon around the earth. The r-squared for the signal is 0.970. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.45 and -0.75. The r-squared associated with this model is 0.026, giving a total r-squared of 0.996, indicating that 99.6% of variation in the data can be explained by this model. The fit is thus extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values are all smaller than 0.05, indicating that there is statistically significant autocorrelation in the filtered series. Also there is a high autocorrelation at lag 28.

Figure 29 shows the analysis for the heights of the first low tide at Songkhla.



signal: $y(t) = 1.001 + 0.2302\cos(at-0.12) + 0.05231\cos(2at+1.339) + 0.03498\cos(11at-0.4136) + 0.1013\cos(13at-0.4661) + 0.09112\cos(14at+2.812) + 0.04537\cos(25at-0.538) + x(t)$
SEs: 0.0004764, 0.0009527

noise: $x(t) = 1.738\,x(t-1) - 0.8291\,x(t-2) + z(t)$ SEs: 0.03063, 0.03071

(t = lunar day, a = 0.01775, r-sq: 0.9404 (signal) + 0.05777 (noise), sd[z] = 0.008963, n = 354)

Figure 29: Analysis of heights of the first low tide at Songkhla in 1996

The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.230, 0.052, 0.035, 0.101, 0.091 and 0.045, respectively. Again the results are similar to those at Pattani. The r-squared for the signal is 0.940. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.74 and -0.83. The r-squared associated with this model is 0.058, giving a total r-squared of 0.998, indicating that 99.8% of variation in the data can explain by this model. The fit extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values are all smaller than 0.05, indicating that there is statistically significant autocorrelation in the filtered series.

Figure 30 shows the analysis for the heights of the second high tide at Songkhla.

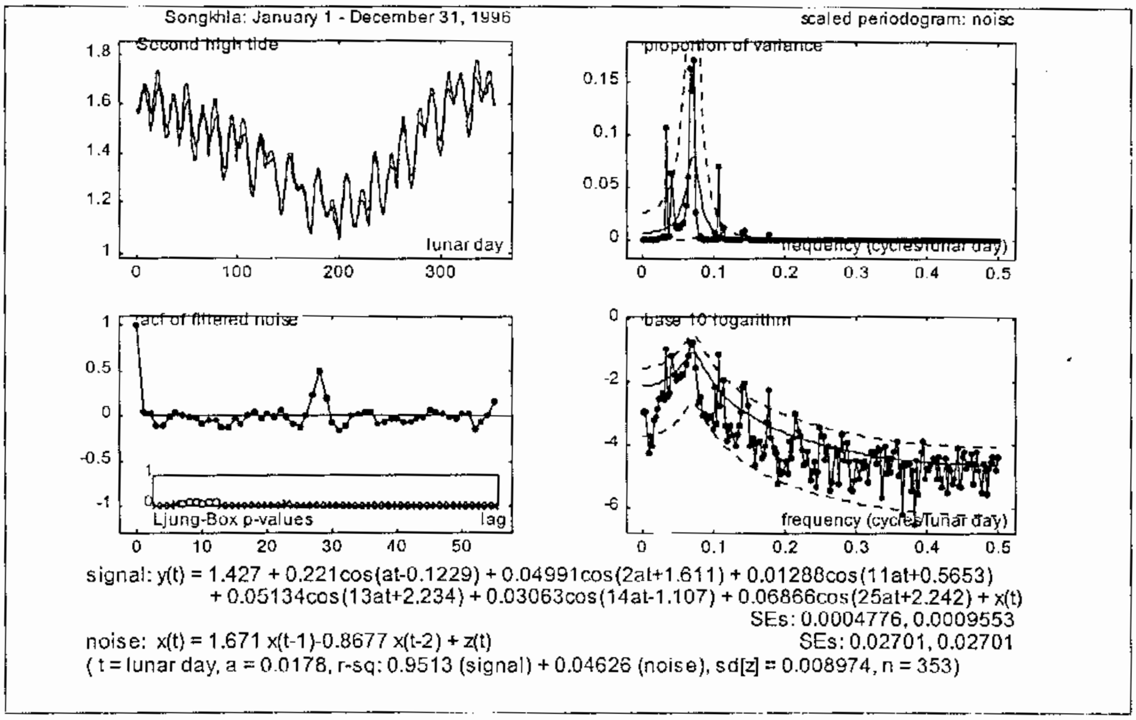


Figure 30: Analysis of heights of the second high tide at Songkhla in 1996

The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.22, 0.05, 0.05 and 0.05, respectively. The r-squared for the signal is 0.951. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.67 and -0.87. The r-squared associated with this model is 0.046, giving a total r-squared of 0.997, indicating that 99.7% of variation in the data can be explained by this model. The fit is thus extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values are greater than 0.05 only for lag 5-12, indicating that the stochastic part of the model is inadequate. Also there is a high autocorrelation at lag 28.

Figure 31 shows the analysis for the heights of the second low tide at Songkhla.

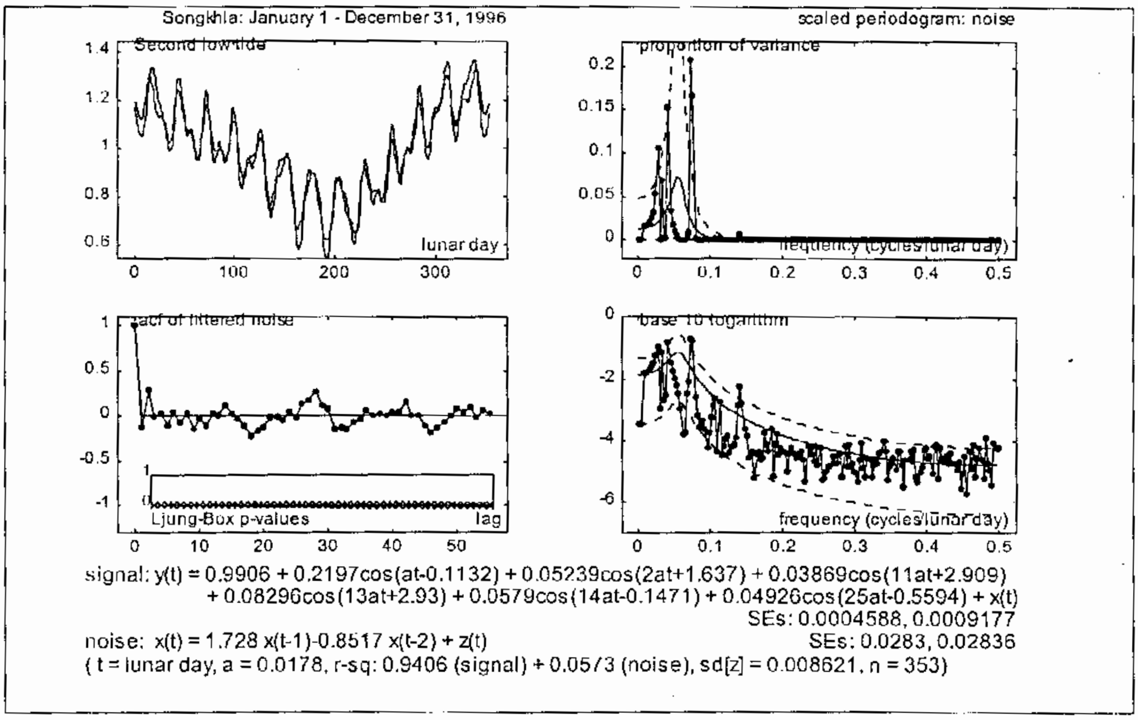


Figure 31: Analysis of heights of the second low tide at Songkhla in 1996

The fitted signal includes harmonics at 1, 2, 11, 13, 14 and 25 cycles with amplitudes 0.220, 0.052, 0.039, 0.083, 0.058 and 0.049, respectively. The r-squared for the signal is 0.940. After subtracting the signal, the noise may be modelled simply as a second order autoregressive process with parameters 1.73 and -0.85. The r-squared associated with this model is 0.057, giving a total r-squared of 0.997, indicating that 99.7% of variation in the data can be explained by this model. The fit is thus extremely good.

From observing the autocorrelation function it may be seen that the Ljung-Box p-values are all smaller than 0.05, indicating that there is statistically significant autocorrelation in the filtered series.

Overall, the results at Pattani and Songkhla are similar, but there are some important differences. These results will be discussed further in Chapter 5.

5. Overall summary of results

5.1 Times of occurrence

signal: $y_j(t) = c_j + A_{11}\cos(11at+\phi_{j11}) + A_{12}\cos(12at+\phi_{j12}) + A_{13}\cos(13at+\phi_{j13})$

$+ A_{14}\cos(14at+\phi_{j14}) + A_{24}\cos(24at+\phi_{j24}) + A_{25}\cos(25at+\phi_{j25})$

$+ A_{26}\cos(26at+\phi_{j26})$; where a= 0.0178 and j is station (Pattani and Songkhla)

Table 1: Signal parameters for times of occurrence

Tides	c	A_{11}	A_{12}	A_{13}	A_{14}	A_{24}	A_{25}	A_{26}	r-sq(signal)
Low 1 (Pattani)	0.092	-	0.372	0.449	-	0.178	0.800	0.267	0.583
Low 1 (Songkhla)	0.186	-	0.242	0.328	-	0.113	0.820	0.143	0.592
High 1 (Pattani)	6.098	-	0.209	0.344	-	0.159	0.809	0.245	0.602
High 1 (Songkhla)	6.462	-	0.079	0.553	-	0.076	0.492	0.048	0.686
Low 2 (Pattani)	12.45	-	0.112	0.588	-	0.196	0.658	0.279	0.714
Low 2 (Songkhla)	12.57	-	0.120	0.563	-	0.159	0.620	0.163	0.713
High 2 (Pattani)	18.41	-	0.122	0.886	-	0.172	0.711	0.279	0.662
High 2 (Songkhla)	18.61	0.358	0.101	0.576	0.277	0.110	0.405	0.091	0.773

noise: $x_j(t) = b_{j1}\,x(t-1) + b_{j2}\,x(t-2) + b_{j3}\,x(t-3)$; where j is station (Pattani or Songkhla)

Table 2: Noise parameters for times of occurrence

Tides	b_1	b_2	b_3	σ	r-sq(noise)	r-sq(total)
Low 1 (Pattani)	0.895	-0.339	-	0.4486	0.212	0.795
Low 1 (Songkhla)	0.960	-0.373	-	0.3742	0.226	0.818
High 1 (Pattani)	0.276	0.197	-	0.2765	0.298	0.900
High 1 (Songkhla)	1.214	-0.399	-	0.1732	0.247	0.933
Low 2 (Pattani)	0.960	-0.373	-	0.1976	0.226	0.940
Low 2 (Songkhla)	1.381	-0.587	-	0.165	0.237	0.950
High 2 (Pattani)	1.214	-0.399	-	0.2979	0.255	0.917
High 2 (Songkhla)	-	1.822	-1.047	0.1628	0.164	0.937

5.2 Heights of tides

signal : $y_j(t) = c_j + A_1 \cos(at+ \phi_{j1}) + A_2 \cos(2at+ \phi_{j2}) + A_{11} \cos(11at+ \phi_{j11})$

$+ A_{13} \cos(13at+ \phi_{j13}) + A_{14} \cos(14at+ \phi_{j14}) + A_{25} \cos(25at+ \phi_{j25})$;

where $a= 0.0178$ and j is station (Pattani and Songkhla)

Table 3: Signal parameters for heights of tides

Tides	c	A_1	A_2	A_{11}	A_{13}	A_{14}	A_{25}	r-sq(signal)
High 1 (Pattani)	0.860	0.204	0.049	0.027	0.060	0.032	0.065	0.962
High 1 (Songkhla)	1.413	0.218	0.048	0.011	0.037	0.015	0.094	0.970
Low 1 (Pattani)	0.524	0.190	0.057	0.023	0.083	0.059	0.058	0.942
Low 1 (Songkhla)	1.001	0.230	0.052	0.035	0.101	0.091	0.045	0.940
High 2 (Pattani)	0.865	0.210	0.051	0.024	0.047	0.028	0.051	0.960
High 2 (Songkhla)	1.427	0.221	0.050	0.013	0.051	0.031	0.069	0.951
Low 2 (Pattani)	0.517	0.184	0.054	0.016	0.048	0.048	0.064	0.934
Low 2 (Songkhla)	0.991	0.220	0.052	0.039	0.083	0.058	0.049	0.941

noise : $x_j(t) = b_{j1}\, x(t-1) + b_{j2}\, x(t-2)$; where j is station (Pattani or Songkhla)

Table 4: Noise parameters for heights of tides

Tides	b_1	b_2	σ	r-sq(noise)	r-sq(total)
High 1 (Pattani)	1.564	-0.770	0.0010	0.035	0.997
High 1 (Songkhla)	1.45	-0.751	0.0118	0.026	0.996
Low 1 (Pattani)	1.663	-0.760	0.0089	0.056	0.998
Low 1 (Songkhla)	1.738	-0.829	0.0090	0.058	0.998
High 2 (Pattani)	1.647	-0.827	0.0084	0.038	0.998
High 2 (Songkhla)	1.671	-0.868	0.0090	0.046	0.997
Low 2 (Pattani)	1.664	-0.810	0.0092	0.063	0.997
Low 2 (Songkhla)	1.728	-0.852	0.0086	0.057	0.998