Inference on Cycles or Periodicities

- The goal is to propose a parametric model for cycles.
- Bayesian Periodogram: marginal *log-likelihood* of the parametric model.
- Connection of this Bayesian periodogram with the standard or *raw* periodogram.
- Example with a CO2 time series. R code.
- Some properties of the classical periodogram.

Basic model for cycles

- Assume we have a time series X_t observed at arbitrary times $t_1, t_2, \ldots, t_n$.
- Also assume that the time series lacks trend or that the trend has been previously removed by one of our detrending techniques (i.e. differencing, lowess, etc.).
- We wish to estimate periodicities in the data.
- Our basic model is a deterministic cyclical term defined by a cosine plus some noise term:

$$x_{t_i} = r\ cos(\omega t_i + \phi) + \epsilon_i$$

- ω defines the fundamental frequency.
- The associated *cycle*, *periodicity* or *wavelength* is

$\lambda = 2\pi/\omega$.

- ϕ denotes the phase $(0 < \phi < 2\pi)$.
- r $(r > 0)$ is the amplitude of the cosine curve.
- As usual the errors ϵ_i are assumed i.i.d with a Normal distribution. $\epsilon_i \sim N(0, \sigma^2)$
- By applying a known trigonometric identity, we can rewrite the model as

$$x_{t_i} = a\ cos(\omega t_i) + b\ sin(\omega t_i) + \epsilon_i$$

- where a and b are model coefficients with $a = r\ cos(\phi)$; $b = -r\ sin(\phi)$.
- both amplitude and phase can be rexpressed in terms of a and b.

- For the amplitude $r = \sqrt{a^2 + b^2}$.
- For the phase $\phi = tan^{-1}(-b/a)$.
- For equally spaced times $(t_1 = 1, t_2 = 2, \ldots, t_n = n)$ and fixed values of a, b and σ^2, the model is equivalent if we add a multiple of 2π to ω. (Why?)
- To avoid such redundancy, take $\omega < 2\pi$.
- Also, notice that if $0 < \omega < \pi$, we obtain the same model representation for the frequency $2\pi - \omega$ by setting $b = -b$.
- Then, we restrict ω to be between 0 and π.

$$0 \leq \omega \leq \pi$$

- With this restriction, the periodocity λ is between 2 and ∞.

- Notice that if ω is given (known) our basic model is a linear regression model of the form:

$$x_{t_i} = f_i^{'}\beta + \epsilon_i$$

- The parameter vector is $\beta = (a, b)$,
- The *regressor vector* is $f_i^{'} = (c_i, s_i)$ where $c_i = cos(\omega t_i)$ and $s_i = sin(\omega t_i)$.

Summary of Bayes results for the Linear Model

- For the linear model, β and σ^2 are essentially location/scale parameters.
- The default non-informative prior for β and σ^2 is:

$$p(\beta, \sigma^2) \propto 1/\sigma^2$$

- With Bayes theorem the posterior distribution is given by

$$p(\beta, \sigma^2 | x, F) \propto f(x|\beta, \sigma^2)(1/\sigma^2)$$

- Under this prior, the posterior distribution for (β, σ^2) is a *Normal-Gamma* distribution.
- Conditional on σ^2, the posterior for β is a p-dimensional Normal with mean b and a covariance matrix $\sigma^2(F'F)^{-1}$

or $\beta \sim N(b, \sigma^2(F'F)^{-1})$.

- The marginal posterior distribution for σ^2 is an Inverse Gamma with shape parameter $n/2$ and scale parameter $R/2$ or $\sigma^2 \sim IG(n/2, R/2)$
- The product of this p-dimensional Normal and the Inverse Gamma defines the Normal/Gamma posterior.
- For the marginal posterior distribution of β we need

$$p(\beta|x, F) = \int p(\beta, \sigma^2|x, F)d\sigma^2$$

- After some algebraic manipulation, it can be shown that

$$p(\beta|x, F) = c(n, p)|F'F|^{1/2}/(1+(\beta-b)'F'F(\beta-b)/ps^2)^{n/2}$$

- Roughly, for n large $p(\beta|x, F) \approx N(b, s^2(F'F)^{-1})$.

- The marginal density of x given F is,

  $$p(x|F) = \int p(x|\beta, \sigma^2)p(\beta, \sigma^2)d\beta d\sigma^2 = c|F^{'}F|^{-1/2}/R^{(n-p)/2}$$

- Due to the sum of squares factorization, we can establish that

  $$p(x|F) \propto |F^{'}F|^{-1/2}(1 - b^{'}F^{'}Fb/(x^{'}x))^{(p-n)/2}$$

- If we think of F as a "parameter", $p(x|F)$ is a likelihood that could be used to produce inferences on F or on quantities that determine F (*marginal likelihood*).
- Under orthogonality of the F matrix, the evaluation of $p(x|F)$ becomes really easy.
- F orthogonal means that $F^{'}F = kI$

- For the cyclical model we consider $p(x|F)$ as $p(x|\omega)$. This defines the *Bayesian Periodogram.*
- Given a fixed value of ω, the basic cyclical model is a linear model with two parameters.
- In the linear model notation, $x = (x_{t_1}, x_{t_2}, \ldots, x_{t_n})'$, $p = 2$, $\beta = (a, b)'$.
- $f_i' = (c_i, s_i)$ is the ith row of F, where $c_i = cos(\omega t_i)$, $s_i = sin(\omega t_i); i = 1, \ldots, n$.
- Lets simply denote $x_{t_i} = x_i$ and define $C = \sum_{i=1}^{n} c_i^2$, $S = \sum_{i=1}^{n} s_i^2$, $K = \sum_{i=1}^{n} c_i s_i$ and $D = SC - K^2$.
- The MLE or LSE of β, $b = (\hat{a}, \hat{b})'$ is given by:
  - $\hat{a} = \frac{S}{D} \left(\sum_{i=1}^{n} x_i c_i\right) - \frac{K}{D} \left(\sum_{i=1}^{n} x_i s_i\right)$

- $\hat{b} = \frac{C}{D} \left(\sum_{i=1}^n x_i c_i\right) - \frac{K}{D} \left(\sum_{i=1}^n x_i s_i\right)$

- If ω is restricted to the values

$$\omega_j = 2\pi \; j/n; \quad j = 1, \ldots n/2$$

we could use the trigonometric identities

- $\sum_{i=1}^n cos(\omega_j i) = \sum_{i=1}^n sin(\omega_j i) = 0$

- $\sum_{i=1}^n cos(\omega_j i) cos(\omega_l i) = \begin{cases} 0, \; j \neq l \\ n, \; j = l = n/2 \\ n/2, \; j = l \neq n/2 \end{cases}$

- $\sum_{i=1}^n sin(\omega_j i) sin(\omega_l i) = \begin{cases} 0, \; j \neq l \\ 0, \; j = l = n/2 \\ n/2, \; j = l \neq n/2 \end{cases}$

- $\sum_{i=1}^{n} cos(\omega_j i) sin(\omega_l i) = 0$ for all j and l

- These identities imply that for the equally spaced case $t_1 = 1, t_2 = 2, \ldots, t_n = n$, $F'F = (n/2)I_{p \times p}$.
- The MLE of b for $\omega_j \neq n/2$ is:
  - $\hat{a} = (2/n) \sum_{i=1}^{n} x_i cos(\omega_j i)$
  - $\hat{b} = (2/n) \sum_{i=1}^{n} x_i sin(\omega_j i)$
- For n large $C/D \approx S/D \approx (2/n)$ and $K/D \approx 0$ when ω is not close to zero.
- Under the same conditions $(F^{'}F)^{-1} \approx (2/n)I_{p \times p}$ and then

$$b^{'}F^{'}Fb \approx (n/2)(\hat{a}^2 + \hat{b}^2)$$

- This implies

$$p(x|F) \propto (1 - (\hat{a}^2 + \hat{b}^2)n/(2x'x))^{(2-n)/2}$$

- This last formula gives us an approximation for the marginal density $p(x|F)$ or in other words an approximation for the Bayesian Periodogram.
- If we define

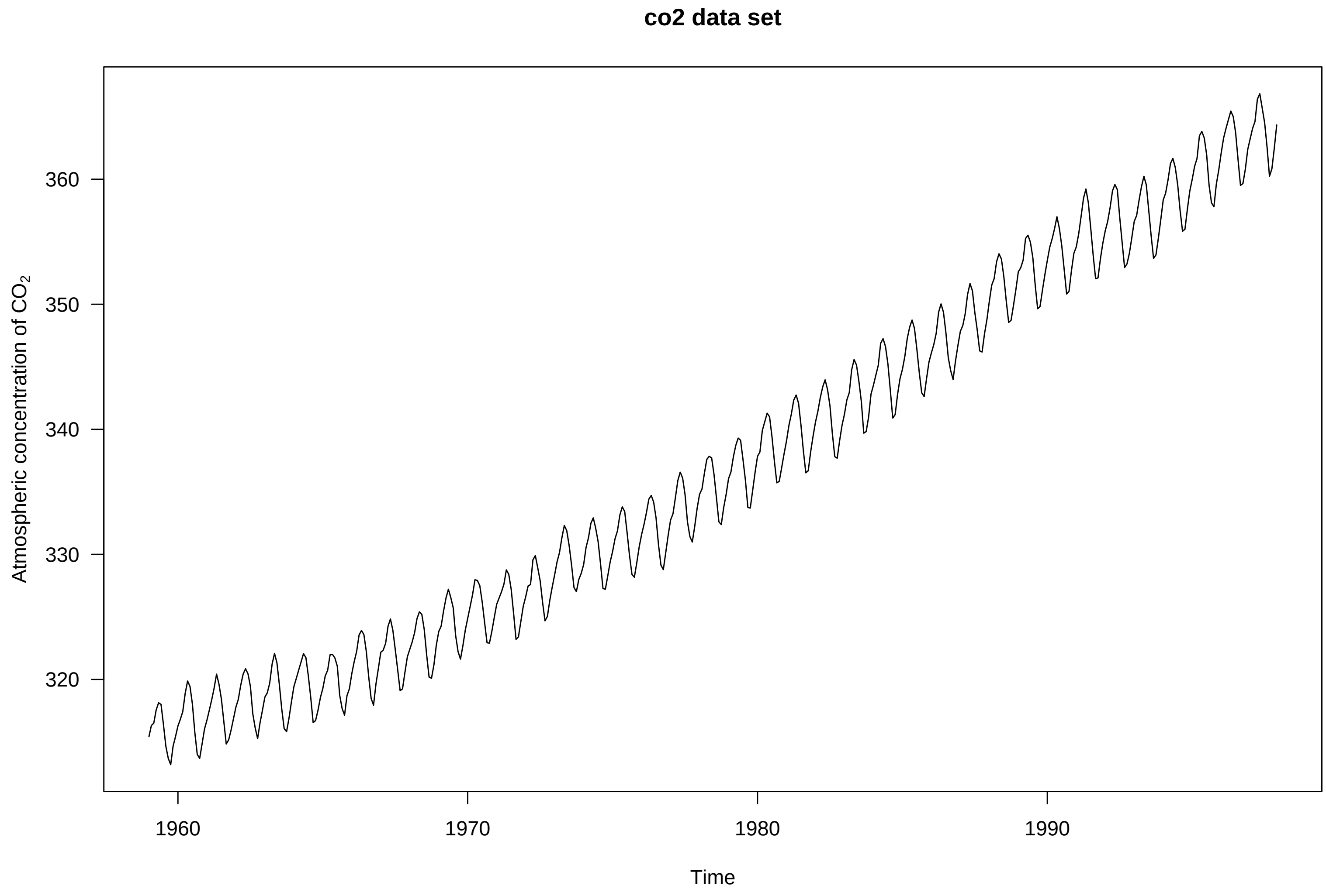
$$I(\omega) = (\hat{a}^2 + \hat{b}^2)/n$$

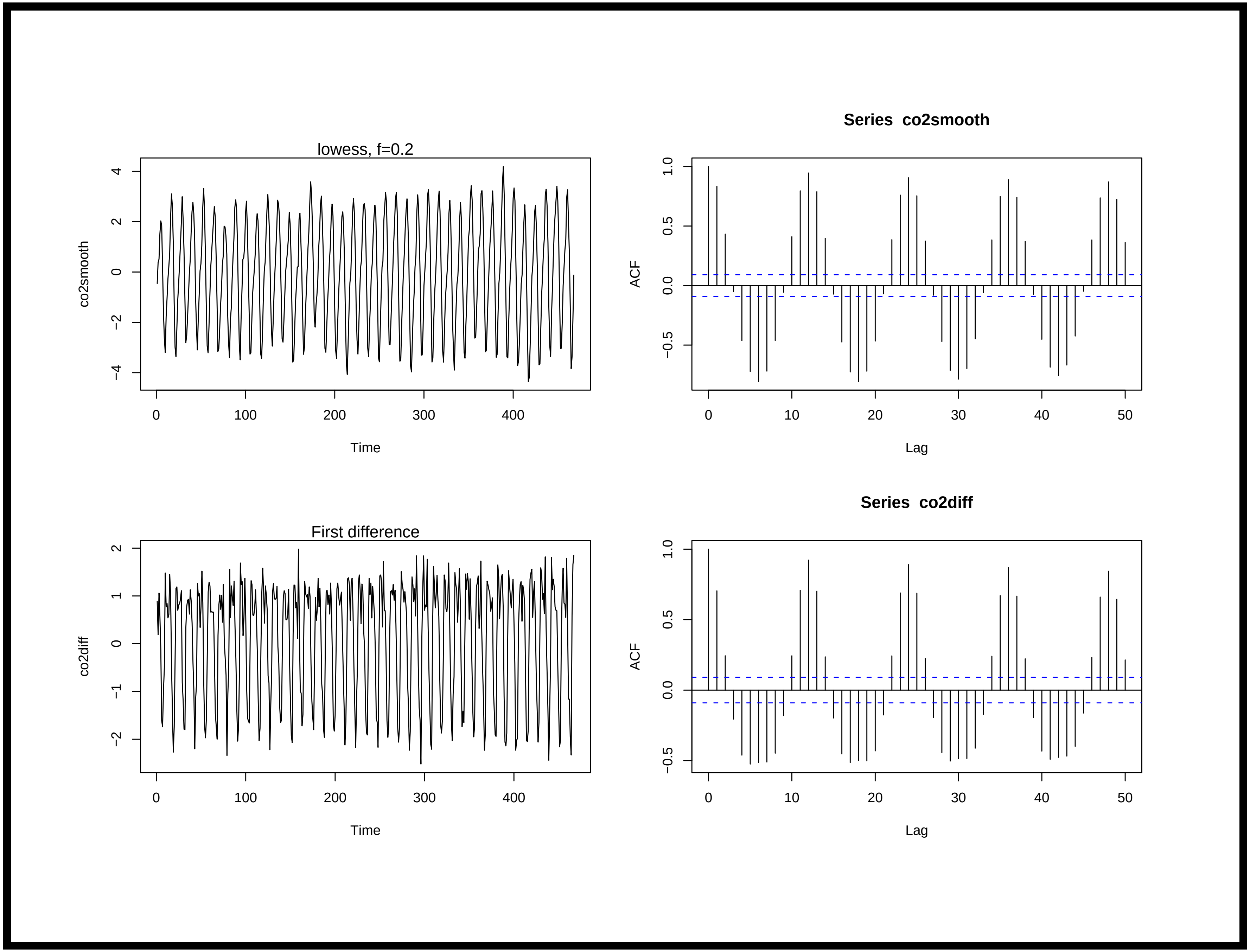
  a plot of ω vs. $I(\omega)$ is known as the *Periodogram*
- $I(\omega)$ is basically the MLE for the amplitude of the sin-cos function that defines our basic model.
- Traditionally, the periodogram is used to find values of ω that produce a high estimated amplitude $I(\omega)$.

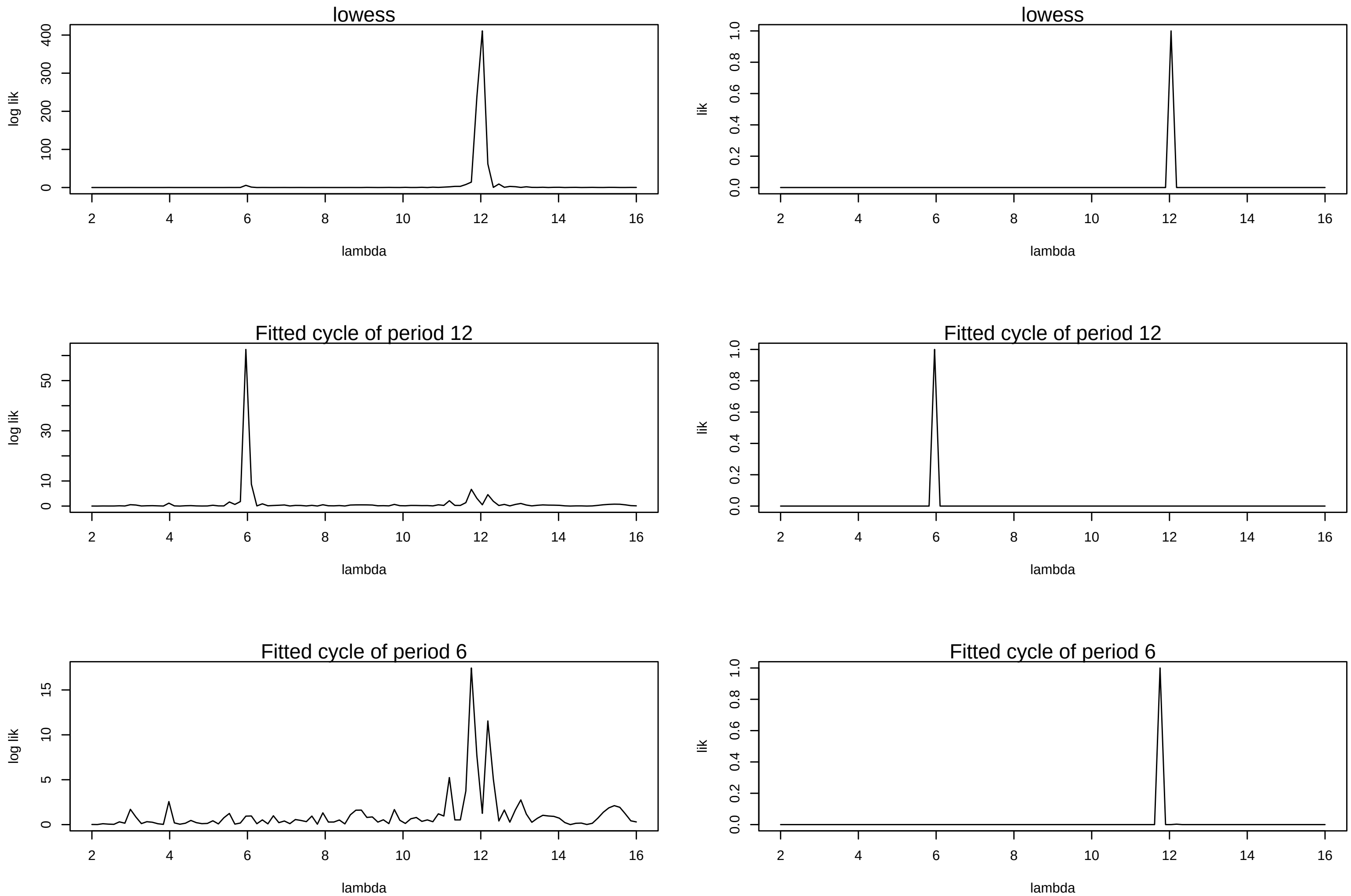
- With the Bayesian Periodogram we look for values of ω that produce a high marginal likelihood $p(x|\omega)$.
- From the approximation, notice that

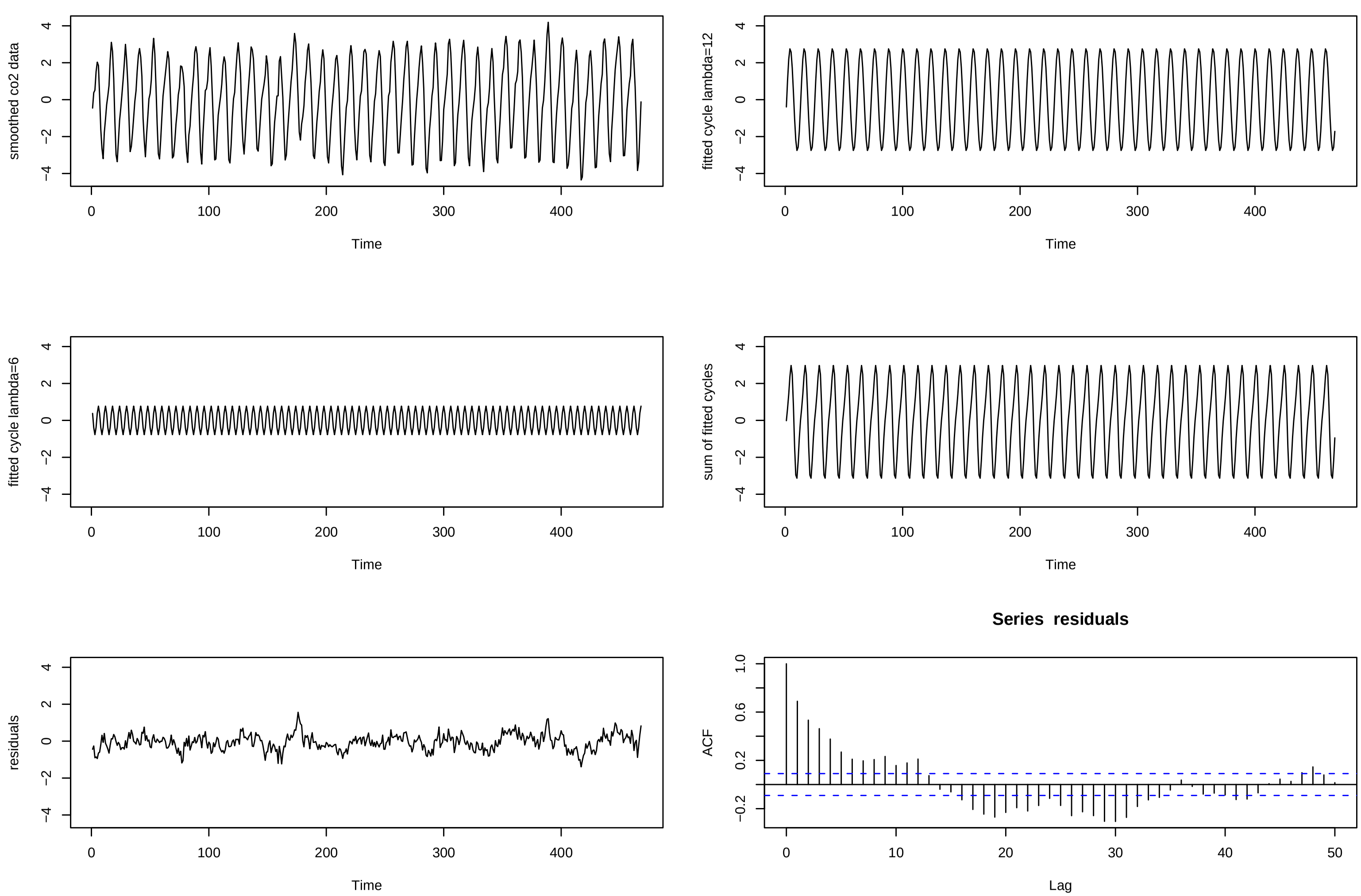
$$log(p(x|F)) \approx (2-n)/2log(1-I(\omega)/(x^{'}x))$$

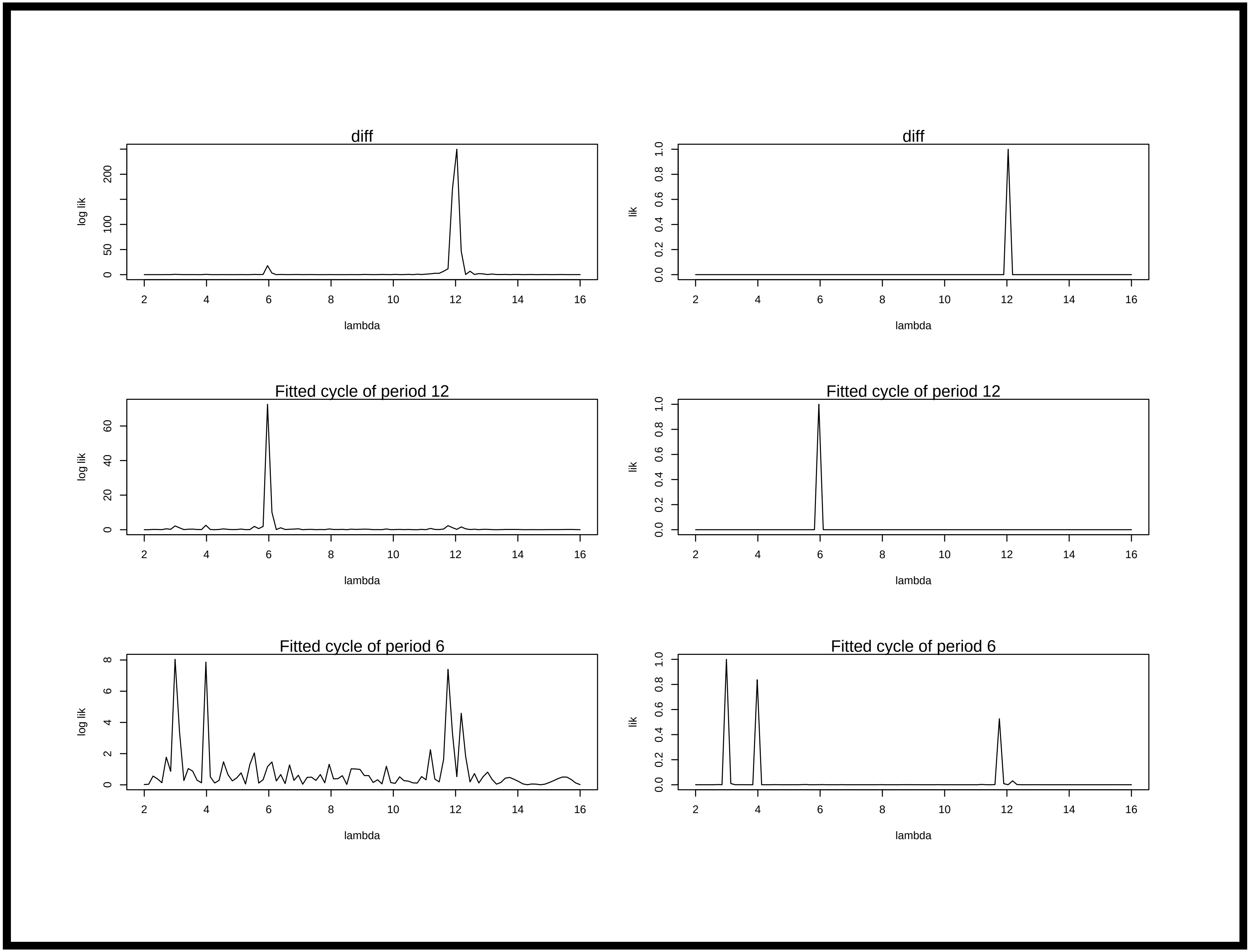
- The Periodogram and the Bayesian Periodogram will contain similar information.
- Illustration with athmospheric concentrations of CO2.
- 468 monthly observations from 1959 to 1997.

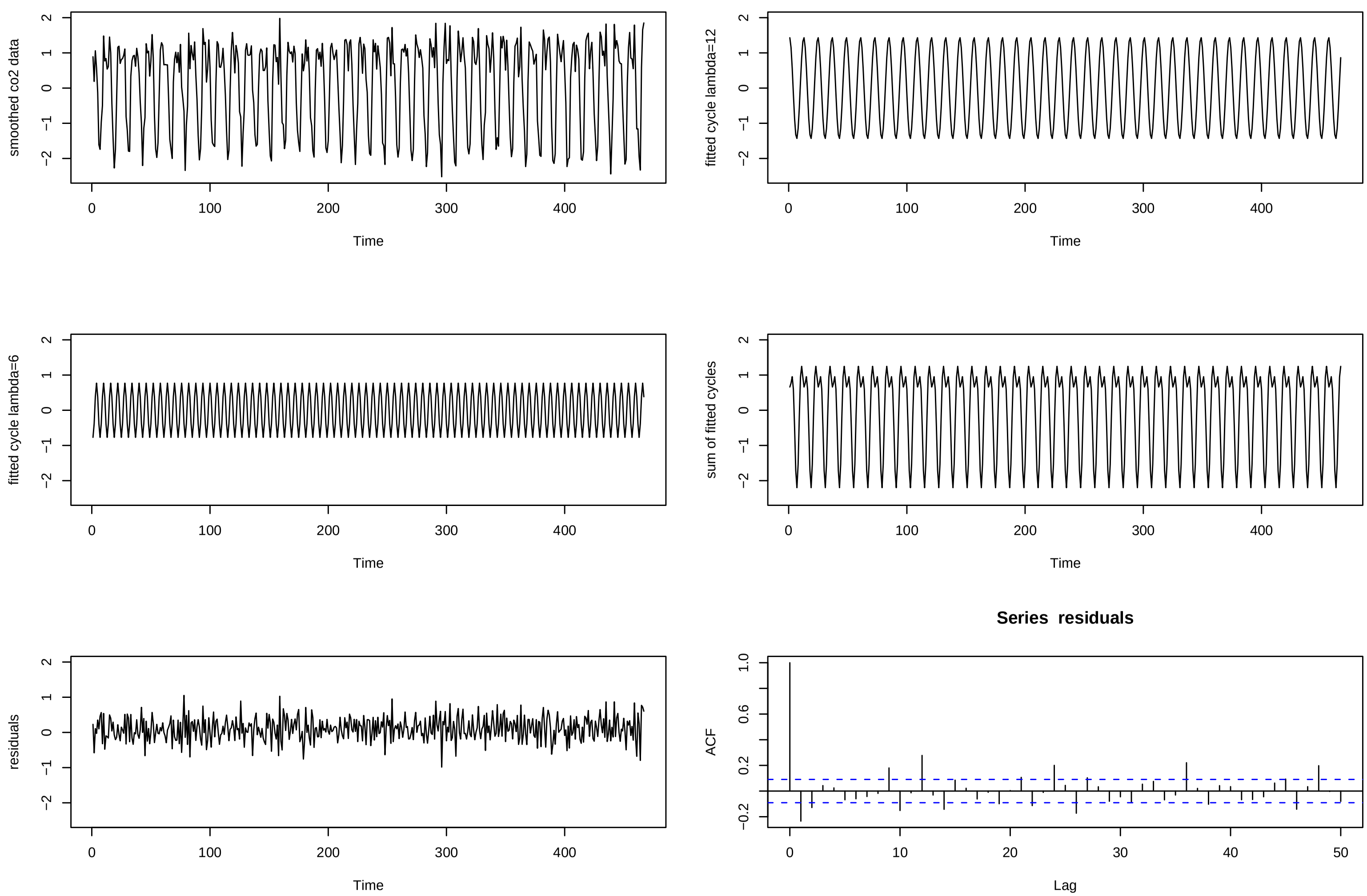

co2 data set
Atmospheric concentration of CO_2
360
350
340
330
320
1960
1970
1980
1990
Time


lowess, f=0.2
co2smooth
Time
Series co2smooth
ACF
Lag
First difference
co2diff
Time
Series co2diff
ACF
Lag


lowess
log lik
lambda
lowess
lik
lambda
Fitted cycle of period 12
log lik
lambda
Fitted cycle of period 12
lik
lambda
Fitted cycle of period 6
log lik
lambda
Fitted cycle of period 6
lik
lambda


smoothed co2 data
fitted cycle lambda=12
fitted cycle lambda=6
sum of fitted cycles
residuals
Series residuals
ACF
Time
Lag


diff
log lik
lambda
diff
lik
lambda
Fitted cycle of period 12
log lik
lambda
Fitted cycle of period 12
lik
lambda
Fitted cycle of period 6
log lik
lambda
Fitted cycle of period 6
lik
lambda


smoothed co2 data
fitted cycle lambda=12
fitted cycle lambda=6
sum of fitted cycles
residuals
Series residuals
ACF
Time
Lag