AMS 223
Midterm
Mickey Warner

Model

We consider a single-component harmonic regression model

$$y_i = a\cos(\omega t_i) + b\sin(\omega t_i) + \epsilon_i, \quad i = 1, \ldots, T$$

where $\epsilon_i \overset{iid}{\sim} N(0, v)$. Let $\boldsymbol{\beta} = (a, b)^\top$ and $\mathbf{f}_i = (\cos(\omega t_i), \sin(\omega t_i))^\top$, then we have $y_i \sim N(\mathbf{f}_i^\top \boldsymbol{\beta}, v)$. Define $\mathbf{y} = (y_1, \ldots, y_T)^\top$ and $\mathbf{F}$ be the $2 \times T$ matrix whose columns are $\mathbf{f}_i$. Then we have

$$\mathbf{y}|\boldsymbol{\beta}, v, \omega \sim N\left(\mathbf{F}^\top \boldsymbol{\beta}, v\mathbf{I}\right).$$

Using the prior $p(\boldsymbol{\beta}, v, \omega) = p(\boldsymbol{\beta}, v|\omega)p(\omega) \propto v^{-1}p(\omega)$, the full conditionals for $\boldsymbol{\beta}$ and v are

$$\begin{aligned}
\boldsymbol{\beta}|v, \omega, \mathbf{y} &\sim N\left(\hat{\boldsymbol{\beta}}, v(\mathbf{F}\mathbf{F}^\top)^{-1}\right) \\
v|\boldsymbol{\beta}, \omega, \mathbf{y} &\sim IG\left(T/2, (\mathbf{y} - \mathbf{F}^\top\boldsymbol{\beta})^\top(\mathbf{y} - \mathbf{F}^\top\boldsymbol{\beta})/2\right),
\end{aligned}$$

where $\hat{\boldsymbol{\beta}} = (\mathbf{F}\mathbf{F}^\top)^{-1}\mathbf{F}\mathbf{y}$ is the maximum likelihood estimate for a fixed ω. For this analysis, interest is primarily in the angular frequency ω (and the corresponding period $\lambda = 2\pi/\omega$). The marginal posterior for ω has the form

$$p(\omega|\mathbf{y}) \propto p(\omega)\left|\mathbf{F}\mathbf{F}^\top\right|^{-1/2}\left[1 - \hat{\boldsymbol{\beta}}^\top\mathbf{F}\mathbf{F}^\top\hat{\boldsymbol{\beta}}/(\mathbf{y}^\top\mathbf{y})\right]^{(2-T)/2}.$$

Direct sampling from the marginal posterior of ω is not avaiable. However, since we're only working with one dimension, we can simply discretize the space for ω, compute $p(\omega|\mathbf{y})$ and sample with replacement based on these probabilities.

We show plots of the frequency and period calculated from the marginal posterior as well as based off samples.

P&W 3.3 – Southern Oscillation Index

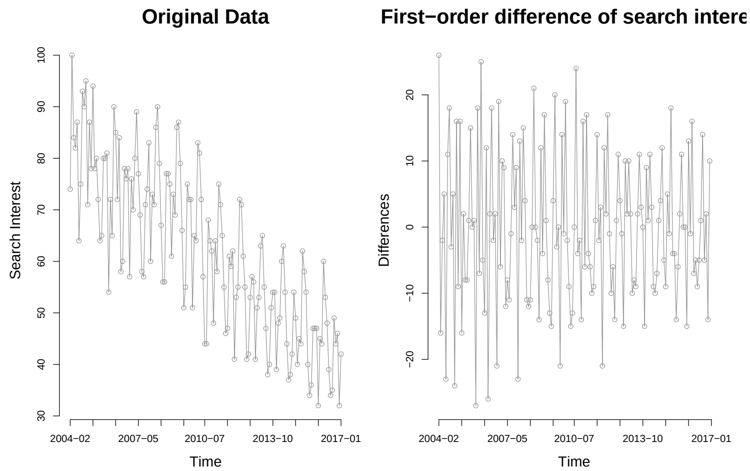

Figure 1: Left: plot of the data. Right: plot of the first differences of the data.

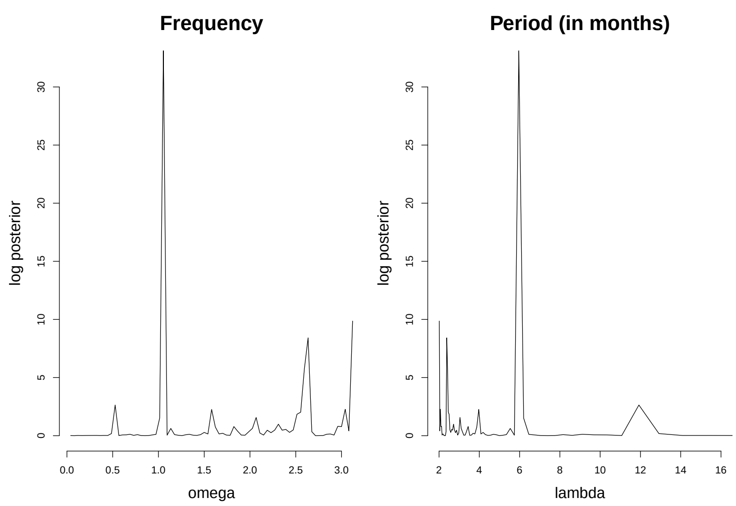

Figure 2: Left: log marginal posterior of the angular frequency $p(\omega|\mathbf{y})$. Right: log marginal posterior of the period/wavelength $p(\lambda = 2\pi/\omega|\mathbf{y})$. The data used is the first differences (right side of Figure 1).

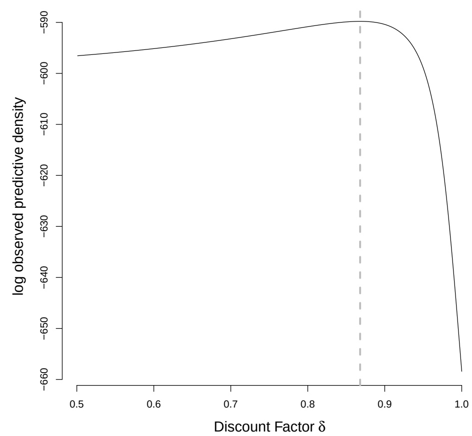

Figure 3: The log observered predictive density evaluated across several $\delta \in [0, 1]$. The dashed line indicates where the maximum occurs, at $\hat{\delta} = 0.868$.

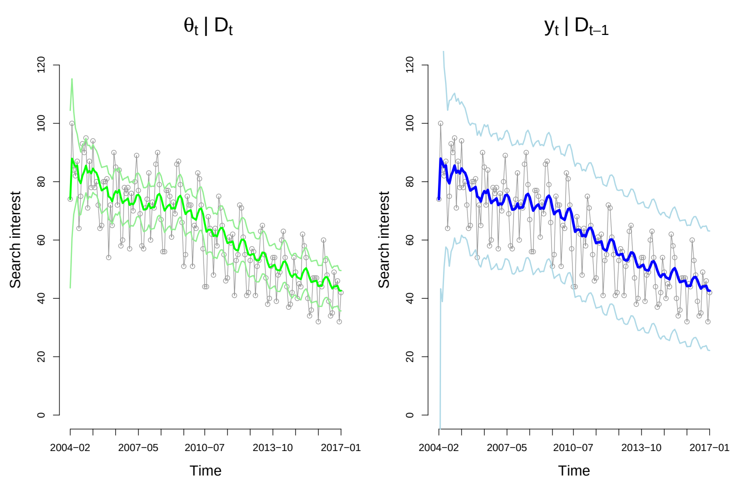

Figure 4: Left: mean of $\theta_t | D_t$ (thick green) with 95% credible intervals (thin green). Right: mean of the one-step ahead forecast function $y_t | D_{t-1}$ (thick blue) with 95% credible intervals (thin blue). Both are evaluated at $t = 1, \ldots, T$.

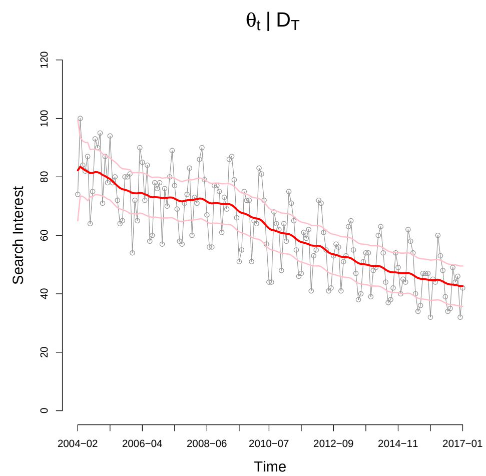

Figure 5: Mean of smoothing distribution $\theta_t|D_T$ (thick red) with 95% credible intervals (thin red) at $t = T-1, \ldots, 1$.

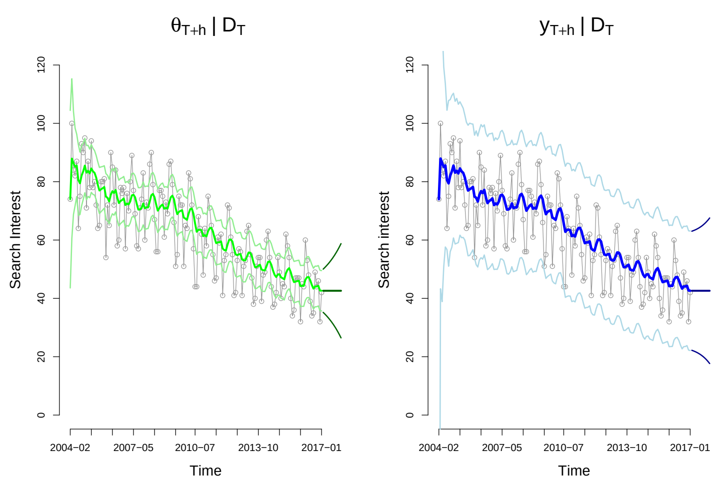

Figure 6: Same as in Figure 4 except the darker shades indicate the forecast distributions $\theta_{T+h}|D_T$ (left) and $y_{T+h}|D_T$ (right) for $h = 1, \ldots, 12$.