AMS 263
Mickey Warner

California precipitation extremes during 1950–1999

Introduction

Methods

Results

Conclusion

The midterm considered the model $\mathcal{M}_0 = \{1, 1, v, vW_t\}$. This model failed to capture any linear trend or seasonal components in the forecast function. We now consider $\mathcal{M}_1 = \{\mathbf{F}, \mathbf{G}, v, v\mathbf{W}_t\}$, where

$$
\begin{aligned}
\mathbf{F} &= (\mathbf{E}_2^\top, \mathbf{E}_2^\top, \mathbf{E}_2^\top, \mathbf{E}_2^\top, \mathbf{E}_2^\top, \mathbf{E}_2^\top, 1)^\top \\
\mathbf{G} &= \text{blockdiag}\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4, \mathbf{G}_5, -1 \right)
\end{aligned}
$$

with

$$
\begin{aligned}
\mathbf{E}_2 &= (1, 0)^\top \\
\mathbf{G}_j &= \begin{pmatrix} cos(2\pi j/p) & sin(2\pi j/p) \\ -sin(2\pi j/p) & cos(2\pi j/p) \end{pmatrix}
\end{aligned}
$$

where $p = 12$ is the fundamental period. The first two elements of $\mathbf{F}$ correspond to the second-order polynomial trend (which results in the first-order trend in the forecast), and the remaining 11 elements are for the seasonal components. The $\mathbf{G}$ matrix is similar. We also let $\mathbf{W}_t$ be specified using a discount factor δ, $\mathbf{W}_t = (1-\delta)/\delta \mathbf{G} \mathbf{C}_{t-1} \mathbf{G}^\top$. $\mathbf{C}_{t-1}$ is the prior scaling matrix for $\boldsymbol{\theta}_{t-1}|\mathcal{D}_{t-1}$.

We use the following priors

$$
\begin{aligned}
\boldsymbol{\theta}_0|\mathcal{D}_0 &\sim T_{n_0}(\mathbf{m}_0, \mathbf{C}_0) \\
v|\mathcal{D}_0 &\sim IG(n_0/2, d_0/2)
\end{aligned}
$$

with $\mathbf{m}_0 = (60, \mathbf{0}_{12}^\top)^\top$, $\mathbf{C}_0 = 200\mathbf{I}_{13}$, $n_0 = 1$, and $d_0 = 100$. These are comparable to the priors I used for $\mathcal{M}_0$ and should be fairly noninformative. Posteriors are obtainined using FFBS.

Analysis

The discount factor δ is selected by maximizing the observed predictive density

$$p_1(\mathbf{y}|\mathcal{D}_0) = \prod_{t=1}^{T} p_1(y_t|\mathcal{D}_{t-1})$$

The subscript 1 denotes we are working under $\mathcal{M}_1$. We evaluate $\log p_1(\mathbf{y}|\mathcal{D}_0)$ on the grid $\delta = \{0.800, 0.801, \ldots, 1.000\}$. The plot of δ vs. $\log p_1(\mathbf{y}|\mathcal{D}_0)$ is given in Figure 1. The vertical line marks where the maximum occurs, at $\hat{\delta}_1 = 0.945$, our selected discount factor. Note, $\log p_1(\mathbf{y}|\mathcal{D}_0) = -478.9$, at $\hat{\delta}_1$.
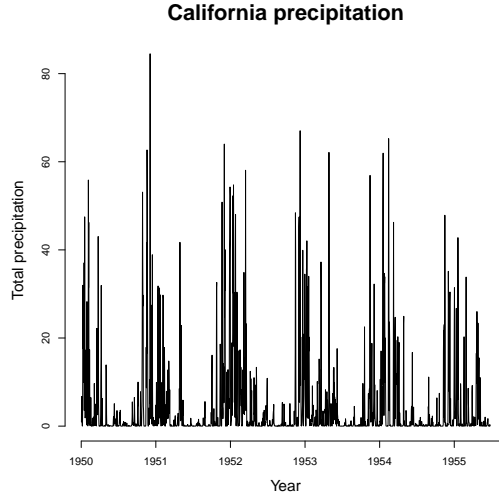

California precipitation

Figure 1: The log observered predictive density evaluated across several $\delta \in [0, 1]$. The dashed line indicates where the maximum occurs, at $\hat{\delta}_1 = 0.945$.

The one-step ahead distribution $y_t|\mathcal{D}_{t-1}$ is summarized in Figure 2. A five-year forecast is shown in Figure 3. The marginal filtering distributions for

the trend and the seasonal components (i.e. $\theta_{t,1}$ and $\theta_{t,3}, \theta_{t,5}, \theta_{t,7}, \theta_{t,9}, \theta_{t,11}$ conditioned on $\mathcal{D}_t$) are given in Figure 4. The smoothing distributions (i.e. conditioned on $\mathcal{D}_T$) are in given in Figure 5.
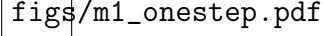

figs/m1_onestep.pdf

Figure 2: The gray dots mark the data. The thick red line is the mean forecast function and the thin dashed lines are 95% credible intervals for $\mathcal{M}_1$.
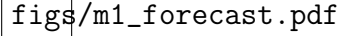

figs/m1_forecast.pdf

Figure 3: The five-year forecast for $\mathcal{M}_1$, for $h = 1, \ldots, 60$. The vertical line marks where the data end, picking up at the end of the plot in Figure 2.
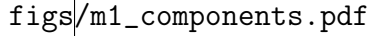

figs/m1_components.pdf

Figure 4: The trend and seasonal components for $\mathcal{M}_1$. The little numbers in the top right for the harmonic plots indicate the proportion of times the upper bound from the 95% interval fell below zero. The little numbers in the bottom right are for the lower bound going above zero.
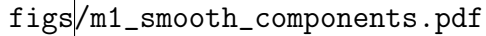

figs/m1_smooth_components.pdf

Figure 5: Same as in Figure 4, but for the smoothing distributions.

The forecast distribution seems to be much more reasonable than what we saw from $\mathcal{M}_0$ which was constant for all $t > T$. The forecast in $\mathcal{M}_1$ continues the downward trend that we're seeing in the data as well as the cyclical behavior. A deficiency in our forecast function (and our model) is that we do not account for the bounds. If this forecast function were to have its way, we would see the search interest for UCSC fall to oblivion, being not worth more than a bare mention.

Consider the distribution of the harmonics given in Figure 5. We see that harmonics 1, 2, and 5, have 95% intervals that cross zero at least 30% of the time (marked by the little numbers on the right side of the plots). Harmonic 4 rarely cross zero, suggesting that it is unimportant since it may as well be a straight line. Harmonic 3 crosses zero about 17% of the time, providing some evidence that this harmonic is unimportant. We discard harmonics 3 and 4 and perform the analysis.

The reduced model has a similar $\mathbf{F}$ and $\mathbf{G}$ except in $\mathbf{F}$ the fourth and fifth $\mathbf{E}_2^\top$ are omitted and $\mathbf{G}_3$ and $\mathbf{G}_4$ are omitted from $\mathbf{G}$. We obtain a discount factor of $\hat{\delta}_2 = 0.952$ which evaluates to a log observed predictive density at $\log p_2(\mathbf{y}|\mathcal{D}_0) = -482.3$, 3.4 less than that of $\mathcal{M}_1$. The summary of distributions in Figure 6 through 9 are highly comparable to those from $\mathcal{M}_1$.

Since the removal of harmonics 3 and 4 result in only a small decrease in the log observed predictive density, we would say that the reduced model $\mathcal{M}_2$ is as useful as $\mathcal{M}_1$.

Figure 6: The gray dots mark the data. The thick red line is the mean forecast function and the thin dashed lines are 95% credible intervals for $\mathcal{M}_2$.

Figure 7: The five-year forecast for $\mathcal{M}_2$, for $h = 1, \ldots, 60$. The vertical line marks where the data end, picking up at the end of the plot in Figure 6.

Figure 8: The trend and seasonal components for $\mathcal{M}_2$. The little numbers in the top right for the harmonic plots indicate the proportion of times the upper bound from the 95% interval fell below zero. The little numbers in the bottom right are for the lower bound going above zero.

Figure 9: Same as in Figure 8, but for the smoothing distributions.

Algorithm for fitting $\mathcal{M}_3$

Model $\mathcal{M}_3$ is as follows:

$$y_t = \alpha + \beta t + x_t + \epsilon_t, \qquad \epsilon_t \sim N(0, v),$$

$$x_t = \sum_{j=1}^{q} \phi_j x_{t-j} + \nu_t, \qquad \nu_t \sim N(0, 30)$$

with $\alpha \sim N(0, vw)$, $\beta \sim N(0, vw)$, and v, w unknown and q known.

The conjugate prior for v is the inverse gamma, $v \sim IG(a_v, b_v)$, as well as for w, $w \sim IG(a_w, b_w)$. We let $x_{1-j} \sim N(m_j, C_j)$, for $j = 1, \ldots, q$ be the priors for $x_0, x_{-1}, \ldots, x_{1-q}$. Since ϕ are (conditional on $\mathbf{x}$) AR coefficients, a good prior to use is $\phi \sim N(\mathbf{m}_\phi, \mathbf{C}_\phi)$, subject to being in the stationary region.

We obtain posterior samples for $(x_{1:T}, \alpha, \beta, \phi, v, w)$ with the following algorithm.

1. Initialize $(x_{1:T}^{(0)}, \phi^{(0)}, v^{(0)}, w^{(0)})$. Set $i = 0$.

2. Increment i by 1.

3. Sample $(\alpha^{(i)}, \beta^{(i)})^\top = \boldsymbol{\mu}^{(i)}$ from
   $\pi(\boldsymbol{\mu}|\cdot) \propto N(\mathbf{y} - \mathbf{x}^{(i-1)}|\mathbf{Z}_T\boldsymbol{\mu}, v^{(i-1)}\mathbf{I}_T)N(\mathbf{0}, v^{(i-1)}w^{(i-1)}\mathbf{I}_2)$, where the rows of $\mathbf{Z}_T$ are $(1, t)$, for $t = 1, \ldots, T$.

4. Sample $v^{(i)}$ from $v|\cdot \sim IG\left(a_v, b_v + \frac{1}{2}(K + (\alpha^{(i)})^2/w^{(i-1)} + (\beta^{(i)})^2/w^{(i-1)})\right)$ where $K = (\mathbf{y} - \mathbf{x}^{(i-1)} - \mathbf{Z}_T\boldsymbol{\mu}^{(i)})^\top(\mathbf{y} - \mathbf{x}^{(i-1)} - \mathbf{Z}_T\boldsymbol{\mu}^{(i)})$.

5. Sample $w^{(i)}$ from $w|\cdot \sim IG\left(a_w, b_w + \frac{1}{2}((\alpha^{(i)})^2/v^{(i)} + (\beta^{(i)})^2/v^{(i)})\right)$.

6. Since, conditional on $\phi^{(i-1)}$, we have a DLM, then we sample from $\mathbf{x}^{(i)}$ using FFBS. Use the recursion formulas to get $\pi(x_t|\mathcal{D}_t, \cdot)$, then obtain the smoothing distributions $\pi(x_t|\mathcal{D}_T, \cdot)$, for $t = T : 1$. Sample $x_t^{(i)}$ from $\pi(x_t|\mathcal{D}_T, \cdot)$, where $(\cdot)$ denotes the most recent samples from the other parameters.

7. Sample $\phi^{(i)}$ using a Metropolis-Hastings step. The function used in the M-H ratio is $\pi(\phi|\mathbf{x}, \cdot) \propto \pi(\phi) \prod_{t=1}^{T} N(x_t|\sum_{i=1}^{j} \phi_j x_{t-j}, 30)$

8. Go to step 2 until $i = B$, for some large B.