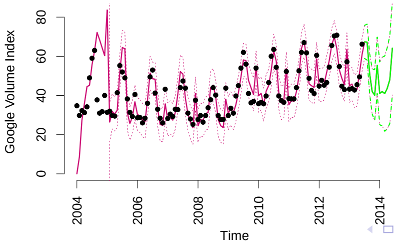
DLM Overview

- Notation.
- Models.
- Bayesian inference.
- Examples.

Why Bayesian dynamic models?

Bayesian dynamic modeling and forecasting comprises:

- sequential model definitions for series of observations over time;
- structuring via models with useful parameterizations;
- a **probabilistic** representation of information about **all** parameters and observables;



Dynamic models: Notation

- *Initial information set:* denoted as $\mathcal{D}_0$. Represents **all** the available information used to form initial views about the future;
- *Information set at time t* : $\mathcal{D}_t$. If no other information is available at time t, $\mathcal{D}_t = \{y_t, \mathcal{D}_{t-1}\}$. If additional information I_t is available $\mathcal{D}_t = \{I_t, \mathcal{D}_{t-1}\}$.
- *Forecast distribution:* $(y_s|\mathcal{D}_t), \quad s > t$.

Dynamic models: Notation

The following elements need to be specified:

- a parametric model at the observational level for t :

$$p(y_t|\boldsymbol{\theta}_t, \phi);$$

- a parametric model describing the evolution of $\boldsymbol{\theta}_t$ over time:

$$p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \phi),$$

- prior distributions $p(\boldsymbol{\theta}_0, \phi|\mathcal{D}_0)$.

Here ϕ denote additional known constants or uncertain quantities.

Dynamic models: Notation

Key goals and quantities:

- **Learning:** Filtering and smoothing densities $p(\theta_t, \phi|\mathcal{D}_t)$ and $p(\theta_t, \phi|\mathcal{D}_T)$, for $T > t$.
- One-step and k-steps ahead **forecast distributions:** $p(y_{t+1}|\mathcal{D}_t)$ and $p(y_{t+k}|\mathcal{D}_t)$.

General univariate **dynamic linear model (DLM)**:

$$
\begin{array}{rrcl}
\text{[Observation]} & y_t & = & \mathbf{F}_t' \boldsymbol{\theta}_t + \nu_t, \\
\text{[System]} & \boldsymbol{\theta}_t & = & \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \\
\text{[Initial info]} & \boldsymbol{\theta}_0 & \sim & p(\boldsymbol{\theta}_0 | \mathcal{D}_0),
\end{array}
$$

with

- $\boldsymbol{\theta}_t = (\theta_{t,1}, \ldots, \theta_{t,p})'$ the state vector.
- $\mathbf{F}_t$ a p-dimensional vector of known constants.
- $\mathbf{G}_t$ the $p \times p$ evolution matrix.
- ν_t the observation noise and $\mathbf{w}_t$ the state evolution noise.
- In the **normal DLM (NDLM),** $\nu_t \sim N(0, v_t)$, $\mathbf{w}_t \sim N(\mathbf{0}, \mathbf{W}_t)$, with ν_s and $\mathbf{w}_t$ independent and mutually independent and $(\boldsymbol{\theta}_0 | \mathcal{D}_0) \sim N(\mathbf{m}_0, \mathbf{C}_0)$.

- Shorthand notation: $\{\mathbf{F}_t, \mathbf{G}_t, v_t, \mathbf{W}_t\}$.
- Information up to time t :

$$\mathcal{D}_t = \{y_t, \mathcal{D}_{t-1}\}.$$

  Usually $\mathcal{D}_t = \{y_{1:t}, \mathcal{D}_0\}$.
- h-step-ahead forecast function for $h \geq 1$:

$$f_t(h) = E(y_{t+h}|\mathcal{D}_t) = \mathbf{F}'_{t+h}\mathbf{G}_{t+h}\ldots\mathbf{G}_{t+1}E(\boldsymbol{\theta}_t|\mathcal{D}_t), \quad h \geq 1$$

Polynomial trend models

- **First order polynomial model** $\{1, 1, v_t, w_t\}$:

$$\begin{aligned} y_t &= \theta_t + \nu_t, \\ \theta_t &= \theta_{t-1} + w_t. \end{aligned}$$

Forecast function: $f_t(h) = a_{t,0}$.

- **Second order polynomial model**

$$\begin{aligned} y_t &= \theta_{t,1} + \nu_t, \\ \theta_{t,1} &= \theta_{t-1,1} + \theta_{t-1,2} + w_{t,1}, \\ \theta_{t,2} &= \theta_{t-1,2} + w_{t,2}. \end{aligned}$$

Then, we have $\{\mathbf{E}_2, \mathbf{J}_2(1), v_t, \mathbf{w}_t\}$, with $\mathbf{E}_2 = (1,0)'$ and

$$\mathbf{J}_2(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Forecast function: $f_t(h) = a_{t,0} + a_{t,1} h$.

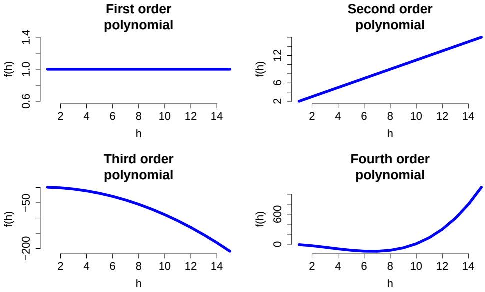
- ***p*-th order polynomial model** (canonical representation) $\{\mathbf{E}_p, \mathbf{J}_p(1), \nu_t, \mathbf{w}_t\}$, with

$$\mathbf{E}_p = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{J}_p(1) = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

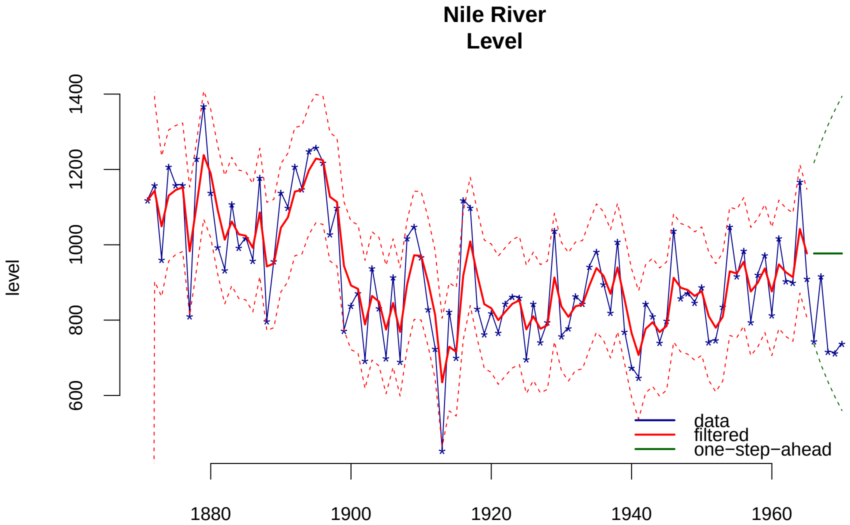
Forecast function: $f_t(h) = a_{t,0} + a_{t,1} h + \cdots + a_{t,p-1} h^{p-1}$.

- Alternative representation for *p*-**th order polynomial DLM:** $\{\mathbf{E}_p, \mathbf{L}_p, \cdot, \cdot\}$ with

$$\mathbf{L}_p = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$



Example: Nile River Data, Petris et al., 2009



- **Seasonal factor models.** Let p be the period and $\boldsymbol{\theta}_t$ the p-dimensional state vector of seasonal levels, with $\theta_{t,1}$ the level at the current time. The seasonal factors model is $\{\mathbf{E}_p, \mathbf{P}, v_t, \mathbf{w}_t\}$ with

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Forecast function: $f_t(h) = \mathbf{E}_p' \mathbf{P}^h \mathbf{m}_t = m_{t,j}$, with $j = p|h$.
Additional constraints are imposed on the seasonal factors. Typically, $\sum_{i=1}^{p} \theta_{t,i} = 0 \Rightarrow$ *form-free seasonal effects model.*

- **Component Fourier representation.** If ω in $(0, \pi)$, the *harmonic component DLM* is $\{\mathbf{E}_2, \mathbf{J}_2(1, \omega), \cdot, \cdot\}$ with

$$\mathbf{J}_2(1, \omega) = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix}.$$

If $\omega = \pi$ we have $\{1, -1, \cdot, \cdot\}$. The *forecast functions* are, respectively,

$$f_t(h) = a_t \cos(\omega h + b_t) \quad \text{for} \quad \omega \in (0, \pi),$$

and

$$f_t(h) = (-1)^h a_t \quad \text{for} \quad \omega = \pi.$$

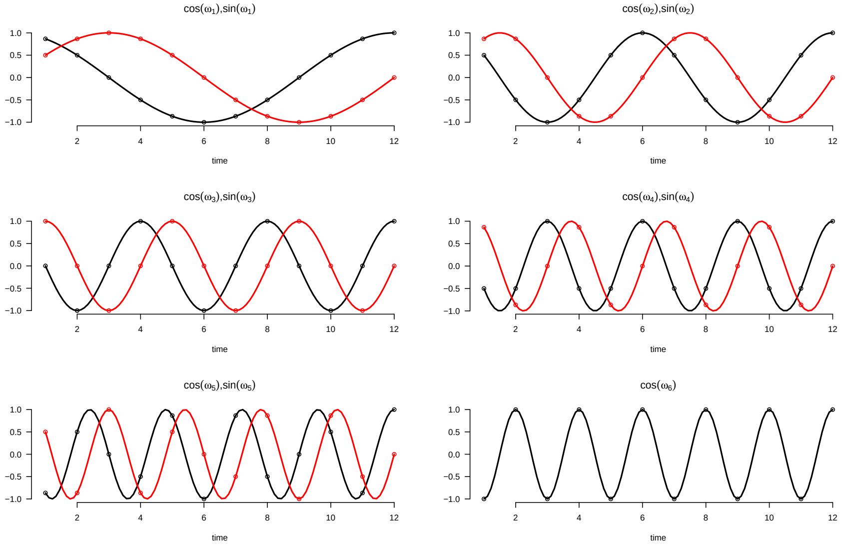
Complete Fourier Representation

- If <u>p is odd</u> $p = 2m - 1$, and the *full effects Fourier form DLM* is $\{\mathbf{F}, \mathbf{G}, \cdot, \cdot\}$ with
  - $\mathbf{F} = (\mathbf{E}_2', \ldots, \mathbf{E}_2')'$,
  - $\mathbf{G} = \text{blockdiag}(\mathbf{G}_1, \ldots, \mathbf{G}_{m-1})$, with $\mathbf{G}_j = \mathbf{J}_2(1, \omega_j)$, $\omega_j = 2\pi j/p$, for $j = 1 : (m-1)$.
- If <u>p is even</u> $p = 2m$, and the model is specified via
  - $\mathbf{F} = (\mathbf{E}_2', \ldots, \mathbf{E}_2', 1)'$
  - $\mathbf{G} = \text{blockdiag}(\mathbf{G}_1, \ldots, \mathbf{G}_{m-1}, -1)$

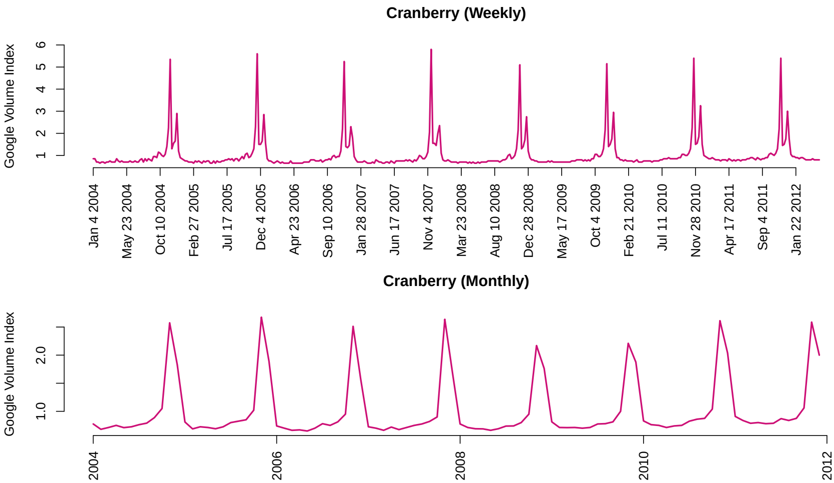
Forecast function:

$$f_t(h) = \sum_{j=1}^{m-1} a_{t,j} \cos(\omega_j h + b_{t,j}) + (-1)^h a_{t,m},$$

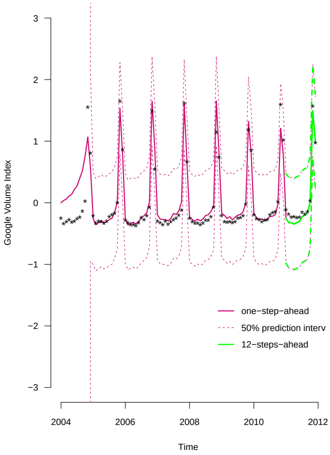
with $a_{t,m} = 0$ if p is odd.

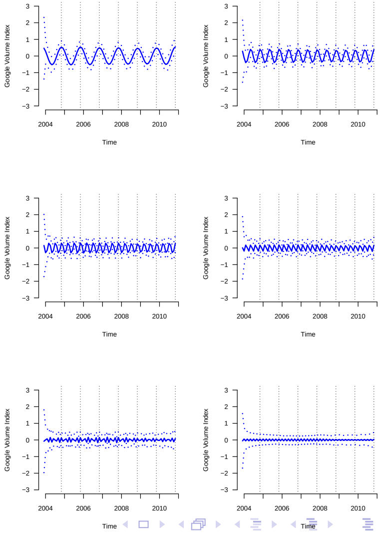

cos(ω1),sin(ω1)
cos(ω2),sin(ω2)
cos(ω3),sin(ω3)
cos(ω4),sin(ω4)
cos(ω5),sin(ω5)
cos(ω6)
1.0
0.5
0.0
−0.5
−1.0
2
4
6
8
10
12
time

Example: Google Trends



Full seasonal model with $p = 12, v = 0.3, w = 0.01$





- **Simple dynamic regression:**

$$
\begin{aligned}
y_t &= \alpha_t + \beta_t x_t + \nu_t, \\
\alpha_t &= \alpha_{t-1} + w_{t,1}, \\
\beta_t &= \beta_{t-1} + w_{t,2}.
\end{aligned}
$$

Forecast function: $f_t(h) = a_{t,0} + a_{t,1} x_{t+h}$.

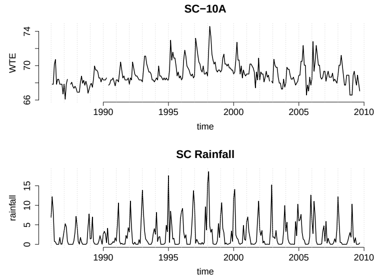
- **Time-varying autoregressions**:

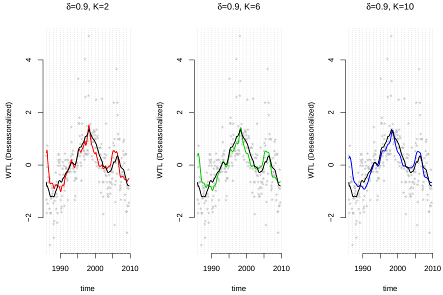
$$
\begin{aligned}
y_t &= \sum_{i=1}^{p} \phi_{t,i} y_{t-i} + \nu_t, \\
\phi_{t,i} &= \phi_{t-1,i} + w_{t,i}, \quad i = 1:p.
\end{aligned}
$$

Forecast function: $f_t(h) = a_{t,1} y_{t+h-1} + \ldots + a_{t,p} y_{t+h-p}$.

- **General regression DLM:** $\{\mathbf{F}_t, \mathbf{G}_t, v_t, \mathbf{W}_t\}$ with $\mathbf{F}_t = (x_{t,1}, \ldots, x_{t,k})'$ and $\mathbf{G}_t = \mathbf{I}_k$.

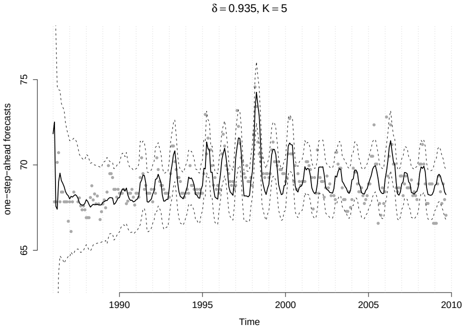
Example: Exploring the relationship between water table elevation data and rainfall at a monitoring well in Santa Cruz, CA. Raw data and anomalies.

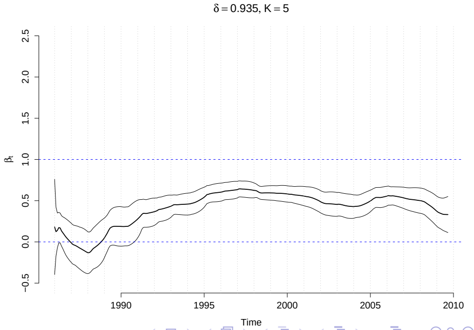




WTE at SC-10A and rainfall

$$\begin{aligned} y_t &= \alpha + \beta_t \times f(\text{rainfall}, t) + \epsilon_t, \\ \beta_t &= \beta_{t-1} + w_t(\delta), \quad \delta \in (0, 1]. \end{aligned}$$



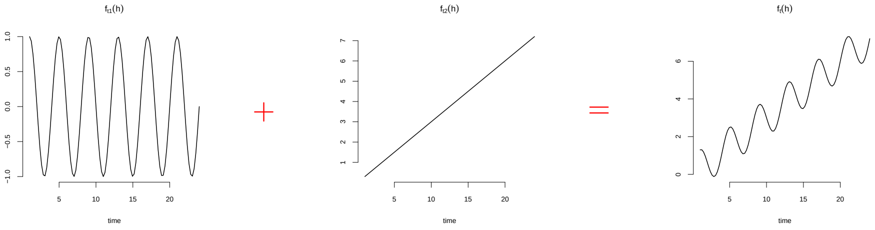


Suppose we have a set of $m > 1$ DLMs $\{\mathbf{F}_{i,t}, \mathbf{G}_{i,t}, v_{i,t}, \mathbf{W}_{i,t}\}$ for $i = 1 : m$. Let $y_t = \sum_{i=1}^{m} y_{i,t}$. Then, y_t has a DLM representation $\{\mathbf{F}_t, \mathbf{G}_t, v_t, \mathbf{W}_t\}$ with

$$\mathbf{F}_t = (\mathbf{F}'_{1,t}, \ldots, \mathbf{F}'_{m,t})', \qquad \mathbf{G}_t = \text{blockdiag}(\mathbf{G}_1, \ldots, \mathbf{G}_m),$$

$$v_t = \sum_{i=1}^{m} v_{i,t}, \qquad \mathbf{W}_t = \text{blockdiag}(\mathbf{W}_{1,t}, \ldots, \mathbf{W}_{m,t}).$$

Forecast function: $f_t(h) = \sum_{i=1}^{m} f_{i,t}(h)$.



Trend + Seasonal Component

- Linear trend model: $\{\mathbf{F}_1, \mathbf{G}_1, \cdot, \cdot\}$ with $\mathbf{F}_1 = (1, 0)'$ and

$$\mathbf{G}_1 = \mathbf{J}_2(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

- Full seasonal model $\{\mathbf{F}_2, \mathbf{G}_2, \cdot, \cdot\}$ with $p = 4$, $\omega = \pi/2$ and with $\mathbf{F}_2 = (1, 0, 1)'$ and

$$\mathbf{G}_2 = \begin{pmatrix} \cos(\pi/2) & \sin(\pi/2) & 0 \\ -\sin(\pi/2) & \cos(\pi/2) & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

$\Rightarrow$

Trend + Seasonal Component

The DLM is a 5-dimensional model $\{\mathbf{F}, \mathbf{G}, \cdot, \cdot\}$ with
$\mathbf{F} = (1, 0, 1, 0, 1)'$ and

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

The forecast function has the form:

$$f_t(h) = (a_{t,1} + a_{t,2}h) + a_{t,3}\cos(\pi h/2) + a_{t,4}\sin(\pi h/2) + a_{t,5}(-1)^h$$

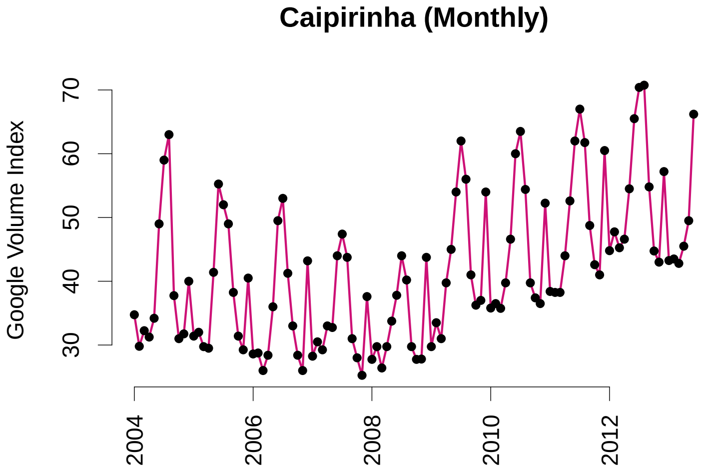
Trend + Regression + Seasonal

Take $p = 12$ and include only the harmonics 1, 3, and 4. The model is $\{\mathbf{F}_t, \mathbf{G}, \cdot, \cdot\}$, with $\mathbf{F}_t' = (1, x_t, \mathbf{E}_2', \mathbf{E}_2', \mathbf{E}_2')'$, and $\mathbf{G} = \text{blockdiag}[1, 1, \mathbf{G}_1, \mathbf{G}_3, \mathbf{G}_4]$, with $\mathbf{E}_2 = (1, 0)'$ and

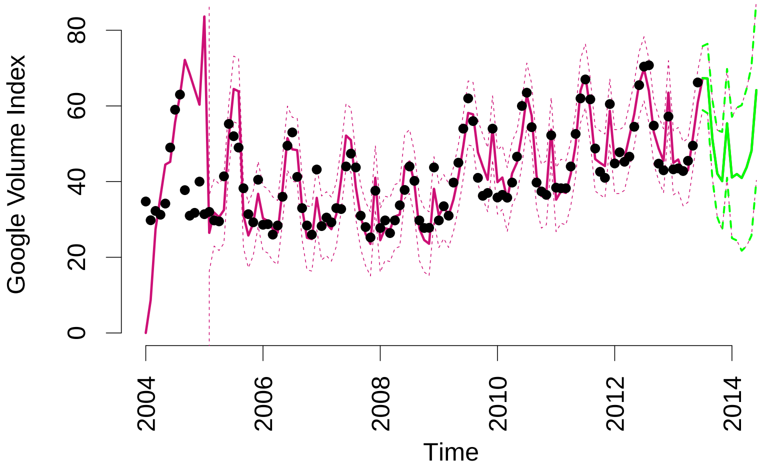
$$\mathbf{G}_r = \begin{pmatrix} \cos(\pi r/6) & \sin(\pi r/6) \\ -\sin(\pi r/6) & \cos(\pi r/6) \end{pmatrix},$$

$r = 1, 3, 4.$

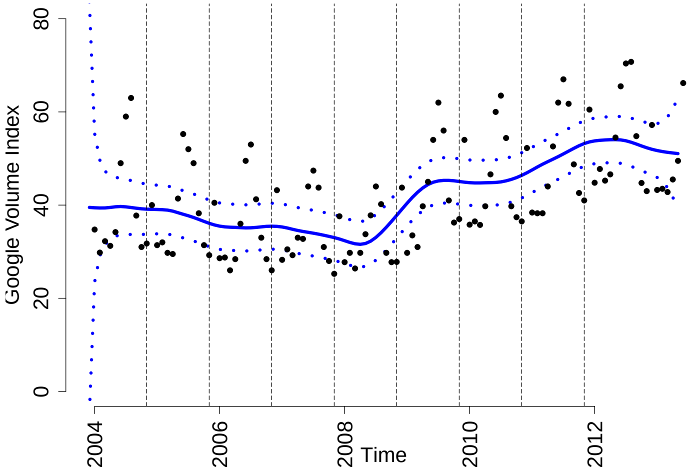
Example: Google Trends



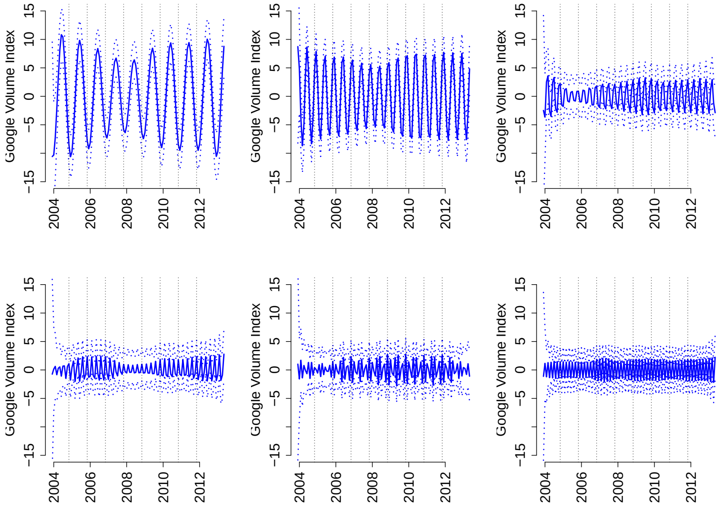
Trend + seasonal, $v = 10, w = 1$: one-step-ahead



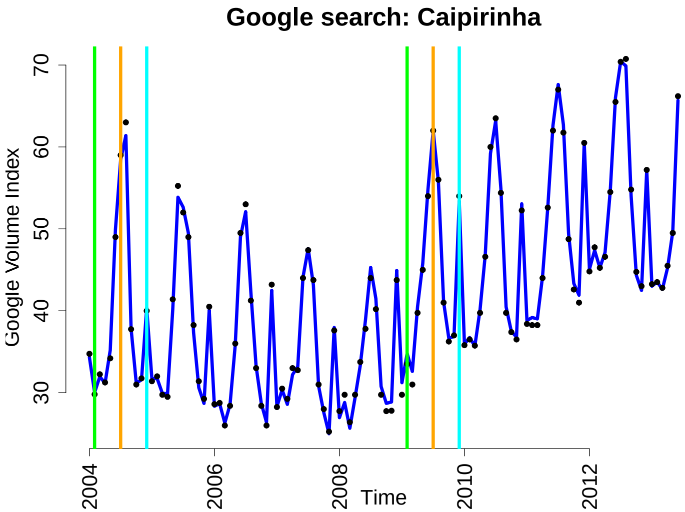
Trend + seasonal, $v = 10, w = 1$: Trend



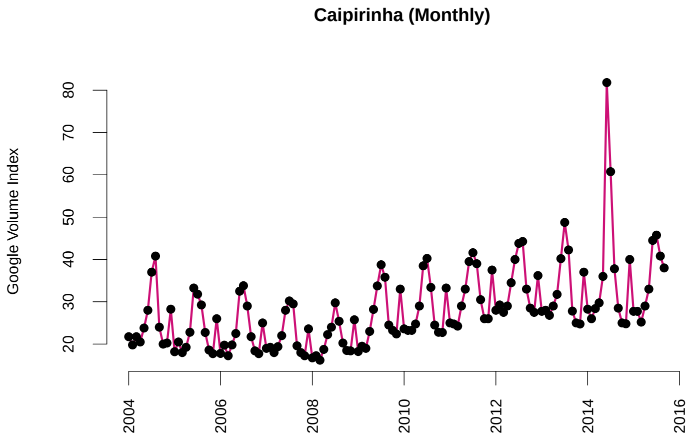
Trend + seasonal, $v = 10, w = 1$: Seasonal



Trend + seasonal, $v = 10, w = 1$: Posterior mean



In case you are curious: Caipirinha (up to 08/2015)...



Learning: v_t and $\mathbf{W}_t$ known

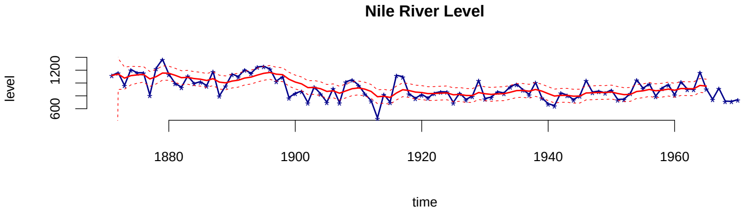
$$\begin{aligned} y_t &= \mathbf{F}_t'\boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N(0, v_t) \\ \boldsymbol{\theta}_t &= \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim N(0, \mathbf{W}_t), \\ (\boldsymbol{\theta}_0|\mathcal{D}_0) &\sim N(\mathbf{m}_0, \mathbf{C}_0). \end{aligned}$$

$$\begin{array}{ccccccccccccc} \boldsymbol{\theta}_0 & \rightarrow & \boldsymbol{\theta}_1 & \rightarrow & \boldsymbol{\theta}_2 & \rightarrow & \cdots & \rightarrow & \boldsymbol{\theta}_{t-1} & \rightarrow & \boldsymbol{\theta}_t & \rightarrow & \boldsymbol{\theta}_{t+1} & \rightarrow \\ & & \uparrow & & \uparrow & & & & \uparrow & & \uparrow & & \uparrow & \\ & & y_1 & & y_2 & & & & y_{t-1} & & y_t & & y_{t+1} & \end{array}$$

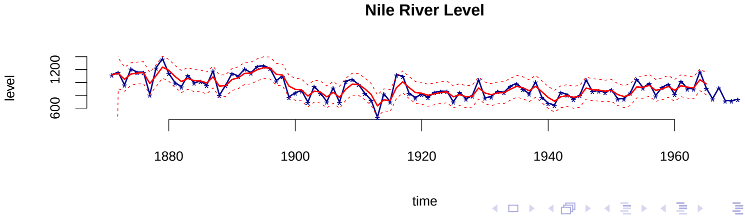
- $(\boldsymbol{\theta}_t|\mathcal{D}_{t-1}) \sim N(\mathbf{a}_t, \mathbf{R}_t)$ with $\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}, \; \mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t' + \mathbf{W}_t$.
- $(y_t|\mathcal{D}_{t-1}) \sim N(y_t|f_t, q_t)$ where $f_t = \mathbf{F}_t'\mathbf{a}_t, \; q_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + v_t$.
  Observing y_t leads to forecast error $e_t = y_t - f_t$.
- $(\boldsymbol{\theta}_t|\mathcal{D}_t) \sim N(\boldsymbol{\theta}_t|\mathbf{m}_t, \mathbf{C}_t)$ with $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t, \; \mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t\mathbf{A}_t' q_t$
  and $\mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t/q_t$.

Nile River Level ([P,P&C, 2.7]). Prior: $m_0 = 0$, $C_0 = 10^7$.

- Model 1: $\{1, 1, 15100, 755\}$, $W/V = 0.05$. $p(\theta_t|\mathcal{D}_t)$:



- Model 2: $\{1, 1, 15100, 7550\}$, $W/V = 0.5$. $p(\theta_t|\mathcal{D}_t)$.



Learning: $v_t = v$ unknown and $\mathbf{W}_t$ known

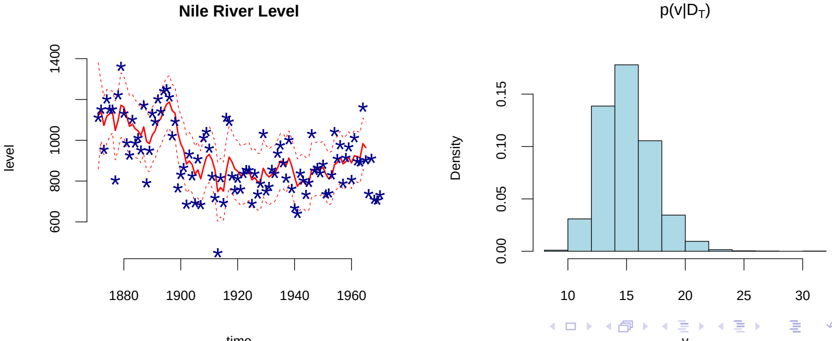
$$
\begin{aligned}
y_t &= \mathbf{F}_t'\theta_t + v_t, \quad v_t \sim N(0, v), \\
\theta_t &= \mathbf{G}_t\theta_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim T_{n_{t-1}}(\mathbf{0}, \mathbf{W}_t), \\
(\theta_0|\mathcal{D}_0) &= T_{n_0}(\mathbf{m}_0, \mathbf{C}_0), \quad (v|\mathcal{D}_0) \sim IG(n_0/2, d_0/2).
\end{aligned}
$$

- $(\theta_t|\mathcal{D}_{t-1}) \sim T_{n_{t-1}}(\mathbf{a}_t, \mathbf{R}_t)$.
- $(y_t|\mathcal{D}_{t-1}) \sim T_{n_{t-1}}(f_t, q_t)$, with $q_t = \mathbf{F}_t'\mathbf{R}_t\mathbf{F}_t + s_{t-1}$ and $s_{t-1} = d_{t-1}/n_{t-1}$.
- $(v|\mathcal{D}_t) \sim IG(n_t/2, n_t s_t/2)$, with $n_t = n_{t-1} + 1$ and $s_t = s_{t-1} + \frac{s_{t-1}}{n_t}\left(\frac{e_t^2}{q_t} - 1\right)$.
- $(\theta_t|\mathcal{D}_t) \sim T_{n_t}(\mathbf{m}_t, \mathbf{C}_t)$, with $\mathbf{C}_t = \frac{s_t}{s_{t-1}}(\mathbf{R}_t - \mathbf{A}_t\mathbf{A}_t' q_t)$.

Example:

Nile River Level ([P,P&C, 4.1])

- Model 3: $\{1, 1, \hat{v}, \hat{w}\}$ with $\hat{v} = 15497.69$ and $\hat{w} = 1213.51$ the MLEs of v and w.
- Model 4: $\{1, 1, v, v\mathbf{W}_t^*\}$. Prior distribution $m_0 = 0$, $C_0^* = 10^4$, $\mathbf{W}_t^* = 100$, $d_0 = 1$ and $n_0 = 10$.



Missing observations

If y_t is missing $\mathcal{D}_t = \mathcal{D}_{t-1}$ and $p(\boldsymbol{\theta}_t|\mathcal{D}_t) = p(\boldsymbol{\theta}_t|\mathcal{D}_{t-1})$. Therefore,

$$\mathbf{m}_t = \mathbf{a}_t, \ \mathbf{C}_t = \mathbf{R}_t, \ n_t = n_{t-1}, \ \text{and } s_t = s_{t-1}.$$

Example: y_3 is missing.

$$\begin{array}{ccccccccccc}
\boldsymbol{\theta}_0 & \rightarrow & \boldsymbol{\theta}_1 & \rightarrow & \boldsymbol{\theta}_2 & \rightarrow & \boldsymbol{\theta}_3 & \rightarrow & \boldsymbol{\theta}_4 & \rightarrow & \cdots \\
 & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
 & & y_1 & & y_2 & & \text{NA} & & y_4 & & \\
 & & & & & & & & & & \\
\mathbf{m}_0 & \mathbf{a}_1 & \mathbf{m}_1 & \mathbf{a}_2 & \mathbf{m}_2 & \mathbf{a}_3 & \mathbf{m}_3 = \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{m}_4 & \mathbf{a}_5 & \cdots \\
\mathbf{C}_0 & \mathbf{R}_1 & \mathbf{C}_1 & \mathbf{R}_2 & \mathbf{C}_2 & \mathbf{R}_3 & \mathbf{C}_3 = \mathbf{R}_3 & \mathbf{R}_4 & \mathbf{C}_4 & \mathbf{R}_5 & \cdots \\
n_0 & & n_1 & & n_2 & & n_3 = n_2 & & n_4 & & \cdots \\
s_0 & & s_1 & & s_2 & & s_3 = s_2 & & s_4 & & \cdots
\end{array}$$

Forecasting

- v_t and $\mathbf{W}_t$ known

$$(\boldsymbol{\theta}_{t+h}|\mathcal{D}_t) \sim N(\mathbf{a}_t(h), \mathbf{R}_t(h)), \text{ and } (y_{t+h}|\mathcal{D}_t) \sim N(f_t(h), q_t(h)),$$

where

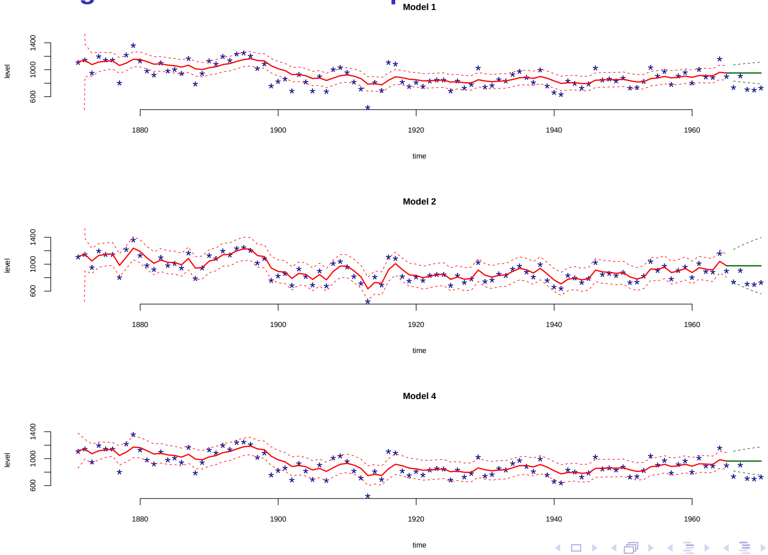
$$\begin{aligned} \mathbf{a}_t(h) &= \mathbf{G}_{t+h}\mathbf{a}_t(h-1), & \mathbf{R}_t(h) &= \mathbf{G}_{t+h}\mathbf{R}_t(h-1)\mathbf{G}'_{t+h} + \mathbf{W}_{t+h}, \\ f_t(h) &= \mathbf{F}'_{t+h}\mathbf{a}_t(h), & q_t(h) &= \mathbf{F}'_{t+h}\mathbf{R}_t(h)\mathbf{F}_{t+h} + v_{t+h}. \end{aligned}$$

- $v_t = v$ unknown and $\mathbf{W}_t$ known

$$(\boldsymbol{\theta}_{t+h}|\mathcal{D}_t) \sim T_{n_t}(\mathbf{a}_t(h), \mathbf{R}_t(h)), \text{ and } (y_{t+h}|\mathcal{D}_t) \sim T_{n_t}(f_t(h), q_t(h)),$$

with $q_t(h) = \mathbf{F}'_{t+h}\mathbf{R}_t(h)\mathbf{F}_{t+h} + s_t$.

Forecasting: Nile River Level Example



Retrospective updating: Smoothing

Let $\mathbf{a}_T(0) = \mathbf{m}_T$, $\mathbf{R}_T(0) = \mathbf{C}_T$ and $t < T$.

- v_t and $\mathbf{W}_t$ known.

$$(\theta_t | \mathcal{D}_T) \sim N(\mathbf{a}_T(t-T), \mathbf{R}_T(t-T)), \quad t = (T-1), (T-2), \ldots$$

with

$$\begin{aligned} \mathbf{a}_T(t-T) &= \mathbf{m}_t - \mathbf{B}_t[\mathbf{a}_{t+1} - \mathbf{a}_T(t-T+1)] \\ \mathbf{R}_T(t-T) &= \mathbf{C}_t - \mathbf{B}_t[\mathbf{R}_{t+1} - \mathbf{R}_T(t-T+1)]\mathbf{B}_t', \end{aligned}$$

where $\mathbf{B}_t = \mathbf{C}_t \mathbf{G}_{t+1}' \mathbf{R}_{t+1}^{-1}$.
In addition, if $\mu_t = \mathbf{F}_t' \theta_t$ is the mean response

$$(\mu_t | \mathcal{D}_T) \sim N(f_T(t-T), \mathbf{F}_t' \mathbf{R}_T(t-T) \mathbf{F}_t),$$

with $f_T(t-T) = \mathbf{F}_t' \mathbf{a}_T(t-T)$.

- $v_t = v$ unknown and $\mathbf{W}_t$ known.

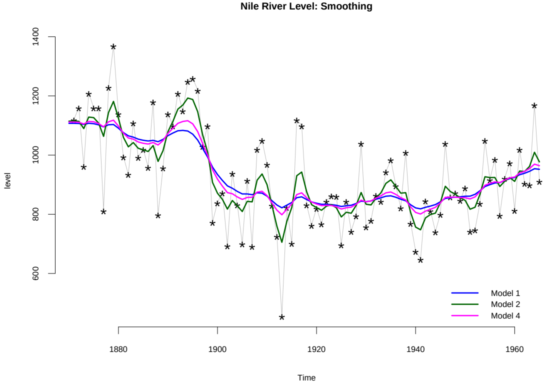
$$(\boldsymbol{\theta}_t|\mathcal{D}_T) \sim T_{n_T}(\mathbf{a}_T(t-T), (s_T/s_t)\mathbf{R}_T(t-T)),$$

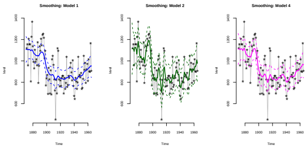
for $t = (T-1), (T-2), \ldots$.

Similarly, we have

$$(\mu_t|\mathcal{D}_T) \sim T_{n_T}(f_T(t-T), (s_T/s_t)\mathbf{F}_t'\mathbf{R}_T(t-T)\mathbf{F}_t).$$

Smoothing: Nile River Level Data





$$\mathbf{R}_t = V(\theta_t | \mathcal{D}_{t-1}) = \mathbf{P}_t + \mathbf{W}_t, \tag{1}$$

where $\mathbf{P}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t'$. $\mathbf{P}_t$: prior variance in a DLM with $\mathbf{W}_t = 0$. Assume that

$$\mathbf{R}_t = \frac{\mathbf{P}_t}{\delta}, \quad \delta \in (0, 1]. \tag{2}$$

Combining (1) and (2) we have that

$$\mathbf{W}_t = \frac{(1-\delta)}{\delta} \mathbf{P}_t.$$

Then, given δ and $\mathbf{C}_0$ $\mathbf{W}_t$ is identified for all t.

- **Choosing** δ. Values in the $(0.8, 1]$ interval are typically relevant in practice. We can choose δ that maximizes

$$\log(\delta) \equiv \log[p(y_{1:T}|\mathcal{D}_0, \delta)] = \sum_{t=1}^{T} \log[p(y_t|\mathcal{D}_{t-1}, \delta)],$$

  or δ that minimizes the MSE or the MAD.
- **Smoothing.** If $\mathbf{G}_t$ is non-singular,

$$\begin{aligned} \mathbf{a}_T(t-T) &= (1-\delta)\mathbf{m}_t + \delta \mathbf{G}_{t+1}^{-1}\mathbf{a}_T(t-T+1) \\ \mathbf{R}_T(t-T) &= (1-\delta)\mathbf{C}_t + \delta^2 \mathbf{G}_{t+1}^{-1}\mathbf{R}_T(t-T+1)(\mathbf{G}_{t+1}')^{-1}. \end{aligned}$$

- **Component discount DLM.** Assume that the models $\{\mathbf{F}_{i,t}, \mathbf{G}_{i,t}, v_{i,t}, \mathbf{W}_{i,t}\}$, for $i = 1 : m$, are superposed. Then, a component discount DLM can be considered by defining $\mathbf{W}_{1,t}, \ldots, \mathbf{W}_{m,t}$ in terms of m discount factors $\delta_1, \ldots, \delta_m$ as

$$\mathbf{W}_{i,t} = \frac{(1-\delta_i)}{\delta_i} \mathbf{P}_{i,t}.$$

Let $\beta \in (0, 1]$ and $\phi_t = 1/v_t$.

$$\begin{aligned} y_t &= \mathbf{F}_t'\boldsymbol{\theta}_t + \nu_t, \;\; \nu_t \sim N(0, v_t), \\ \boldsymbol{\theta}_t &= \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \;\; \mathbf{w}_t \sim T_{n_{t-1}}(\mathbf{0}, \mathbf{W}_t), \\ \phi_t &= \gamma_t\phi_{t-1}/\beta \\ \gamma_t &\sim \text{Beta}(\beta n_{t-1}/2, (1-\beta)n_{t-1}/2). \end{aligned}$$

- $(\boldsymbol{\theta}_t|\mathcal{D}_{t-1}) \sim T_{n_{t-1}}(\mathbf{a}_t, \mathbf{R}_t)$, $(\phi_t|\mathcal{D}_{t-1}) \sim G(\beta n_{t-1}/2, \beta d_{t-1}/2)$ and $(y_t|\mathcal{D}_{t-1}) \sim T_{\beta n_{t-1}}(f_t, q_t)$.
- $(\boldsymbol{\theta}_t|\mathcal{D}_t) \sim T_{n_t}(\mathbf{m}_t, \mathbf{C}_t)$ and $(\phi_t|\mathcal{D}_t) \sim G(n_t/2, d_t/2)$ with $\mathbf{C}_t = (s_t/s_{t-1})(\mathbf{R}_t - \mathbf{A}_t\mathbf{A}_t'q_t)$, $n_t = \beta n_{t-1} + 1$, and

$$d_t = \beta d_{t-1} + s_{t-1}e_t^2/q_t.$$

Ex. Google Cranberry Data.

Model: $\{\mathbf{F}_t, \mathbf{G}_t, v, \mathbf{W}_t\}$ with
$\mathbf{F}_t' = (1, \mathbf{E}_2', \mathbf{E}_2', \mathbf{E}_2') = (1, 1, 0, 1, 0, 1, 0)'$, and
$\mathbf{G}_t = \text{blockdiag}(1, \mathbf{G}_t^s)$, with

$$\mathbf{G}_t^s = \begin{pmatrix} \mathbf{J}_2(1, 2\pi/12) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2(1, 2\pi/6) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_2(1, 2\pi/4) \end{pmatrix},$$

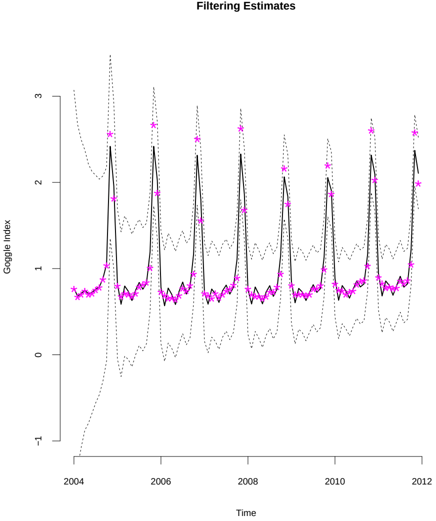
and

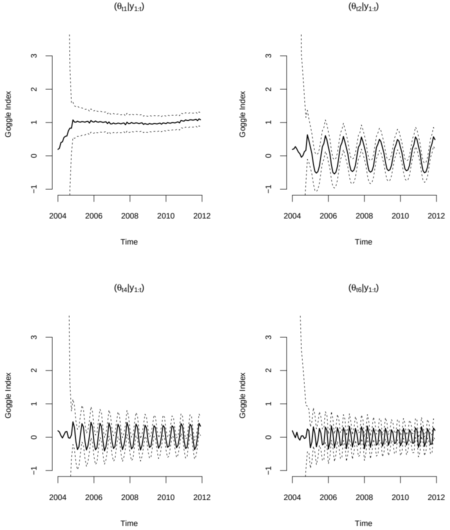
$$\mathbf{J}_2(1, \omega) = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix}.$$

$\mathbf{W}_t$ will be specified as follows:

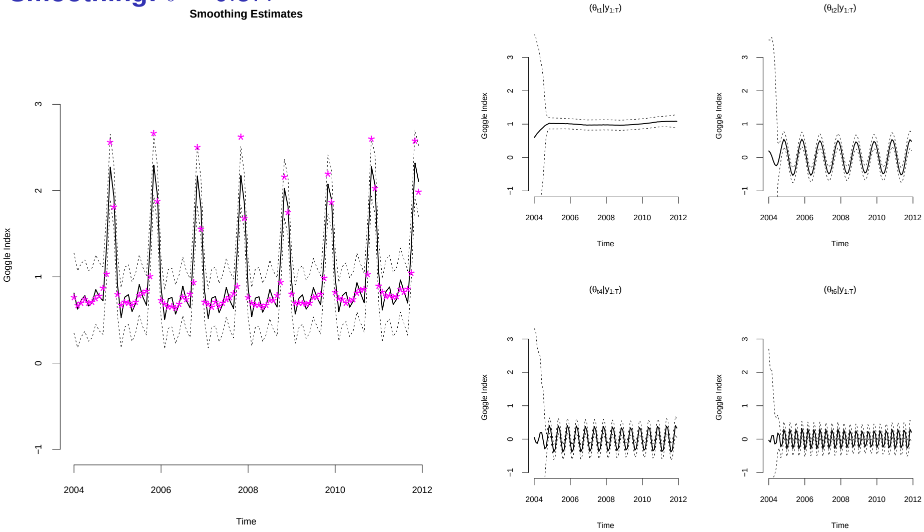
- Using a single discount factor δ.
- Using two discount factors, δ_1 and δ_2, one for the trend and another for the seasonal components.

Filtering: $\delta = 0.87$.

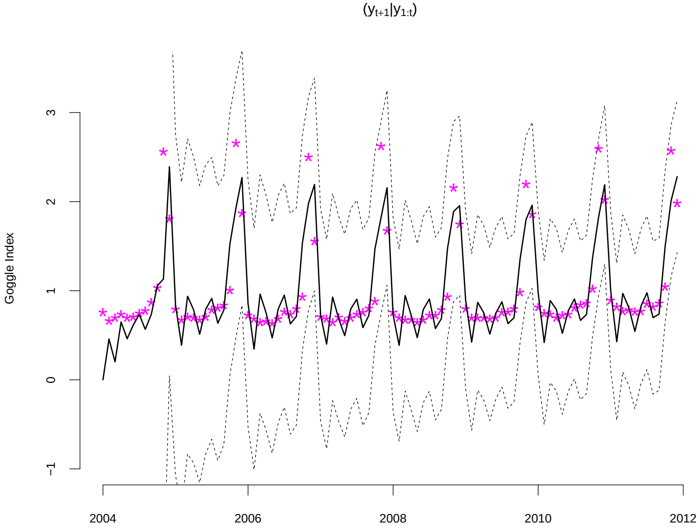




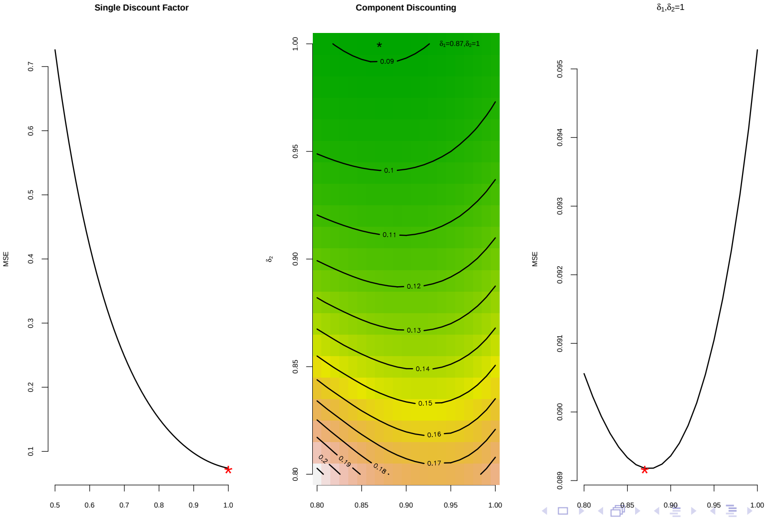
Smoothing: $\delta = 0.87$.

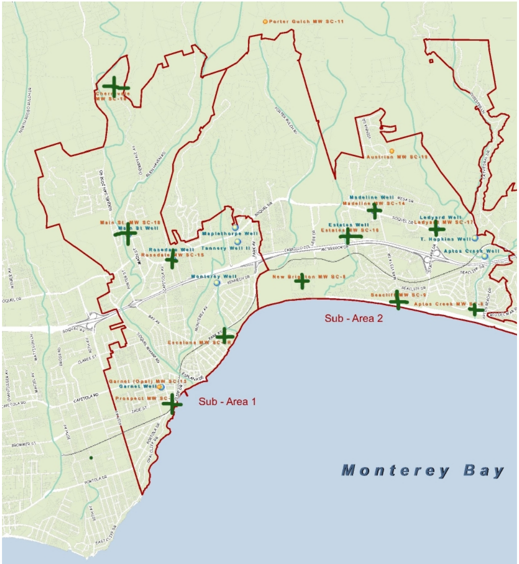


One-Step-Ahead Forecasts: $\delta = 0.87$.



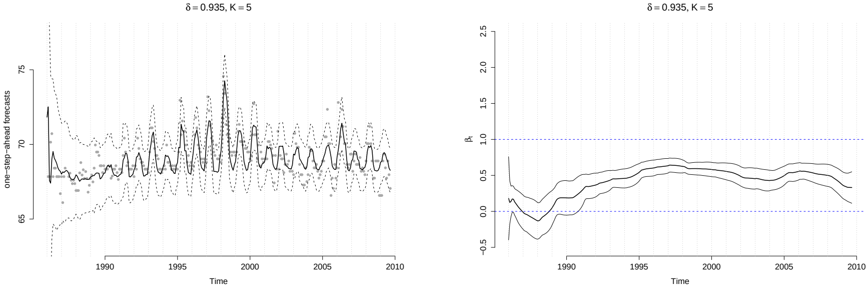
Discount Factors




Porter Gulch MW SC-11
Austrian MW SC-19
Madeline Well
Madeline MW SC-14
Ledyard Well
Estates Well
Estates MW SC-16
T. Hopkins Well
Aptos Creek Well
Main St. MW SC-18
Main St Well
Maplethorpe Well
Tannery Well II
Rosedale Well
Rosedale MW SC-15
Monterey Well
New Brighton MW SC-8
Seacliff MW SC-9
Aptos Creek MW
Sub - Area 2
Garnet (Opal) MW SC-13
Garnet Well
Prospect MW SC
Sub - Area 1
Monterey Bay

WTE at SC-10A and rainfall

$$
\begin{aligned}
y_t &= \alpha + \beta_t \times f(\text{rainfall}_t) + \epsilon_t, \quad \epsilon_t \sim N(0, v) \\
\beta_t &= \beta_{t-1} + w_t(\delta), \quad \delta \in (0, 1] \\
f(\text{rainfall}_t) &= \frac{\sum_{j=0}^{5} \text{rainfall}_{t-j}}{6}.
\end{aligned}
$$



WTE anomalies at SC-10A and rainfall anomalies

$$
\begin{aligned}
y_t^* &= \alpha^* + \beta_t^* \times f^*(\text{rainfall anomalies}_t) + \epsilon_t^*, \quad \epsilon_t^* \sim N(0, v) \\
\beta_t^* &= \beta_{t-1}^* + w_t^*(\delta^*), \quad \delta^* \in (0, 1].
\end{aligned}
$$