Prediction of ESTSP Competition Time Series by Unscented Kalman Filter and RTS Smoother

Simo Särkkä *<simo.sarkka@hut.fi>*
Aki Vehtari *<aki.vehtari@hut.fi>*
Jouko Lampinen *<jouko.lampinen@hut.fi>*

Laboratory of Computational Engineering
Helsinki University of Technology, Finland

February 7, 2007

Contents

Contents

Contents

Contents

State Space Model

- State space model with state $\mathbf{x}_k$ and measurements $\mathbf{y}_k$:

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{q}_{k-1})$$
$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{r}_k),$$

  where $\mathbf{q}_{k-1}$ is the process noise and $\mathbf{r}_k$ is the measurement noise.
- The state $\mathbf{x}_k$ is the hidden internal dynamic state of the system on the time step k
- The measurements $\mathbf{y}_k$ model the output of the system
- We want to estimate the state from the measurements and use it for model based prediction of the time series

State Space Model

- State space model with state $\mathbf{x}_k$ and measurements $\mathbf{y}_k$:

$$\begin{aligned}\mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{q}_{k-1}) \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{r}_k),\end{aligned}$$

  where $\mathbf{q}_{k-1}$ is the process noise and $\mathbf{r}_k$ is the measurement noise.
- The state $\mathbf{x}_k$ is the hidden internal dynamic state of the system on the time step k
- The measurements $\mathbf{y}_k$ model the output of the system
- We want to estimate the state from the measurements and use it for model based prediction of the time series

State Space Model

- State space model with state $\mathbf{x}_k$ and measurements $\mathbf{y}_k$:

$$
\begin{aligned}
\mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{q}_{k-1}) \\
\mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{r}_k),
\end{aligned}
$$

  where $\mathbf{q}_{k-1}$ is the process noise and $\mathbf{r}_k$ is the measurement noise.
- The state $\mathbf{x}_k$ is the hidden internal dynamic state of the system on the time step k
- The measurements $\mathbf{y}_k$ model the output of the system
- We want to estimate the state from the measurements and use it for model based prediction of the time series

State Space Model

- State space model with state $\mathbf{x}_k$ and measurements $\mathbf{y}_k$:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{q}_{k-1}) \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{r}_k), \end{aligned}$$

  where $\mathbf{q}_{k-1}$ is the process noise and $\mathbf{r}_k$ is the measurement noise.
- The state $\mathbf{x}_k$ is the hidden internal dynamic state of the system on the time step k
- The measurements $\mathbf{y}_k$ model the output of the system
- We want to estimate the state from the measurements and use it for model based prediction of the time series

Optimal Filters and Predictors

- Given measurements $\mathbf{y}_1, \ldots, \mathbf{y}_T$ optimal filter produces MMSE optimal online estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

for each $k = 1, \ldots, T$.

- Can be also used for computing the optimal predictions:

$$\hat{\mathbf{x}}(t) = E(\mathbf{x}(t) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

for $t > t_k$.

- If the state space model is linear (i.e., Gaussian process), then Kalman filter provides the optimal solution
- If it is non-linear, then unscented Kalman filter can be used for approximating the optimal solution

Optimal Filters and Predictors

- Given measurements $\mathbf{y}_1, \ldots, \mathbf{y}_T$ optimal filter produces MMSE optimal online estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

  for each $k = 1, \ldots, T$.
- Can be also used for computing the optimal predictions:

$$\hat{\mathbf{x}}(t) = E(\mathbf{x}(t) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

  for $t > t_k$.
- If the state space model is linear (i.e., Gaussian process), then Kalman filter provides the optimal solution
- If it is non-linear, then unscented Kalman filter can be used for approximating the optimal solution

Optimal Filters and Predictors

- Given measurements $\mathbf{y}_1, \ldots, \mathbf{y}_T$ optimal filter produces MMSE optimal online estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

  for each $k = 1, \ldots, T$.

- Can be also used for computing the optimal predictions:

$$\hat{\mathbf{x}}(t) = E(\mathbf{x}(t) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

  for $t > t_k$.

- If the state space model is linear (i.e., Gaussian process), then Kalman filter provides the optimal solution
- If it is non-linear, then unscented Kalman filter can be used for approximating the optimal solution

Optimal Filters and Predictors

- Given measurements $\mathbf{y}_1, \ldots, \mathbf{y}_T$ optimal filter produces MMSE optimal online estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

  for each $k = 1, \ldots, T$.
- Can be also used for computing the optimal predictions:

$$\hat{\mathbf{x}}(t) = E(\mathbf{x}(t) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_k).$$

  for $t > t_k$.
- If the state space model is linear (i.e., Gaussian process), then Kalman filter provides the optimal solution
- If it is non-linear, then unscented Kalman filter can be used for approximating the optimal solution

Optimal Smoothers

- Optimal smoother produces the optimal batch estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_T).$$

- If the dynamic model is linear, then Rauch-Tung-Striebel (RTS) smoother provides the optimal solution
- Approximate smoothers for nonlinear problems exists also
- In this article, the dynamic model is linear (i.e., Gaussian process) given the parameters and thus the linear RTS smoother suffices

Optimal Smoothers

- Optimal smoother produces the optimal batch estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \mid \mathbf{y}_1, \ldots, \mathbf{y}_T).$$

- If the dynamic model is linear, then Rauch-Tung-Striebel (RTS) smoother provides the optimal solution
- Approximate smoothers for nonlinear problems exists also
- In this article, the dynamic model is linear (i.e., Gaussian process) given the parameters and thus the linear RTS smoother suffices

Optimal Smoothers

- Optimal smoother produces the optimal batch estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_T).$$

- If the dynamic model is linear, then Rauch-Tung-Striebel (RTS) smoother provides the optimal solution
- Approximate smoothers for nonlinear problems exists also
- In this article, the dynamic model is linear (i.e., Gaussian process) given the parameters and thus the linear RTS smoother suffices

Optimal Smoothers

- Optimal smoother produces the optimal batch estimate:

$$\hat{\mathbf{x}}(t_k) = E(\mathbf{x}(t_k) \,|\, \mathbf{y}_1, \ldots, \mathbf{y}_T).$$

- If the dynamic model is linear, then Rauch-Tung-Striebel (RTS) smoother provides the optimal solution
- Approximate smoothers for nonlinear problems exists also
- In this article, the dynamic model is linear (i.e., Gaussian process) given the parameters and thus the linear RTS smoother suffices

Idea of the Approach

- **Model** the time series as consisting of periodic and bias components with a linear state space model
- **Model** the signal-residual dependence by including a non-linear correction term into the model
- **Model** the remaining residual autocorrelation with an autoregressive (AR) model
- **Estimate** the parameters and **predict the time series** with the model using the unscented Kalman filter (UKF) and Rauch-Tung-Striebel (RTS) smoother as the numerical methods

Idea of the Approach

- **Model** the time series as consisting of periodic and bias components with a linear state space model
- **Model** the signal-residual dependence by including a non-linear correction term into the model
- **Model** the remaining residual autocorrelation with an autoregressive (AR) model
- **Estimate** the parameters and **predict the time series** with the model using the unscented Kalman filter (UKF) and Rauch-Tung-Striebel (RTS) smoother as the numerical methods

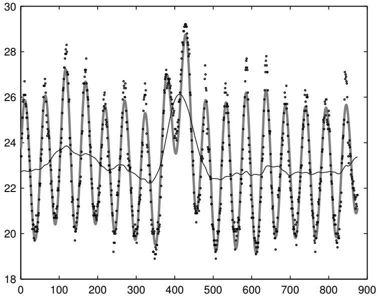
Idea of the Approach

- **Model** the time series as consisting of periodic and bias components with a linear state space model
- **Model** the signal-residual dependence by including a non-linear correction term into the model
- **Model** the remaining residual autocorrelation with an autoregressive (AR) model
- **Estimate** the parameters and **predict the time series** with the model using the unscented Kalman filter (UKF) and Rauch-Tung-Striebel (RTS) smoother as the numerical methods

Idea of the Approach

- **Model** the time series as consisting of periodic and bias components with a linear state space model
- **Model** the signal-residual dependence by including a non-linear correction term into the model
- **Model** the remaining residual autocorrelation with an autoregressive (AR) model
- **Estimate** the parameters and **predict the time series** with the model using the unscented Kalman filter (UKF) and Rauch-Tung-Striebel (RTS) smoother as the numerical methods

Bias and Periodic Components 1/3



- **Signal model**:

$$y_k = x_b(k) + x_r(k) + r_k$$

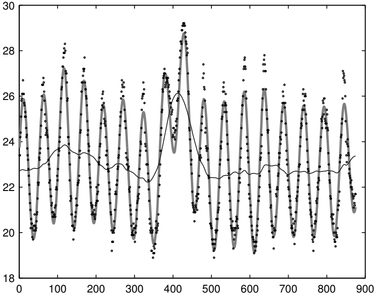
- **Bias** component $x_b(t)$ is modeled as integral of a white noise process $w_b(t)$

$$\frac{\mathrm{d}x_b}{\mathrm{d}t} = w_b(t)$$

- **Periodic** component $x_r(t)$ is modeled as a white noise driven stochastic resonator

$$\frac{\mathrm{d}^2 x_r}{\mathrm{d}t^2} = -\omega^2 x_r + w_r(t)$$

Bias and Periodic Components 1/3



- **Signal model**:

$$y_k = x_b(k) + x_r(k) + r_k$$

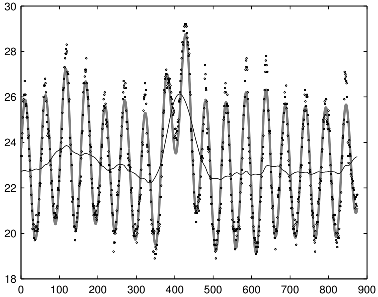
- **Bias** component $x_b(t)$ is modeled as integral of a white noise process $w_b(t)$

$$\frac{\mathrm{d}x_b}{\mathrm{d}t} = w_b(t)$$

- **Periodic** component $x_r(t)$ is modeled as a white noise driven stochastic resonator

$$\frac{\mathrm{d}^2 x_r}{\mathrm{d}t^2} = -\omega^2\, x_r + w_r(t)$$

Bias and Periodic Components 1/3



- Signal model:

$$y_k = x_b(k) + x_r(k) + r_k$$

- **Bias** component $x_b(t)$ is modeled as integral of a white noise process $w_b(t)$

$$\frac{\mathrm{d}x_b}{\mathrm{d}t} = w_b(t)$$

- **Periodic** component $x_r(t)$ is modeled as a white noise driven stochastic resonator

$$\frac{\mathrm{d}^2 x_r}{\mathrm{d}t^2} = -\omega^2 \, x_r + w_r(t)$$

Bias and Periodic Components 2/3

- Can be written as discretely measured continuous-time vector process $\mathbf{x}(t) = (x_b(t)\ x_r(t)\ dx_r(t)/dt)^T$ as follows:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t)$$

- Linear system theory $\Rightarrow$ equivalent discrete time model:

$$\mathbf{x}_k = \mathbf{A}\,\mathbf{x}_{k-1} + \mathbf{q}_{k-1}$$

- Measurement model is of the form

$$y_k = \mathbf{H}\,\mathbf{x}_k + r_k$$

- Linear state space model $\Rightarrow$ **Kalman filter** can be applied.

Bias and Periodic Components 2/3

- Can be written as discretely measured continuous-time vector process $\mathbf{x}(t) = (x_b(t)\ x_r(t)\ dx_r(t)/dt)^T$ as follows:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t)$$

- Linear system theory $\Rightarrow$ equivalent discrete time model:

$$\mathbf{x}_k = \mathbf{A}\,\mathbf{x}_{k-1} + \mathbf{q}_{k-1}$$

- Measurement model is of the form

$$y_k = \mathbf{H}\,\mathbf{x}_k + r_k$$

- Linear state space model $\Rightarrow$ **Kalman filter** can be applied.

Bias and Periodic Components 2/3

- Can be written as discretely measured continuous-time vector process $\mathbf{x}(t) = (x_b(t)\ x_r(t)\ dx_r(t)/dt)^T$ as follows:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t)$$

- Linear system theory $\Rightarrow$ equivalent discrete time model:

$$\mathbf{x}_k = \mathbf{A}\,\mathbf{x}_{k-1} + \mathbf{q}_{k-1}$$

- Measurement model is of the form

$$y_k = \mathbf{H}\,\mathbf{x}_k + r_k$$

- Linear state space model $\Rightarrow$ **Kalman filter** can be applied.

Bias and Periodic Components 2/3

- Can be written as discretely measured continuous-time vector process $\mathbf{x}(t) = (x_b(t)\ x_r(t)\ dx_r(t)/dt)^T$ as follows:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{F}\,\mathbf{x}(t) + \mathbf{L}\,\mathbf{w}(t)$$

- Linear system theory $\Rightarrow$ equivalent discrete time model:

$$\mathbf{x}_k = \mathbf{A}\,\mathbf{x}_{k-1} + \mathbf{q}_{k-1}$$

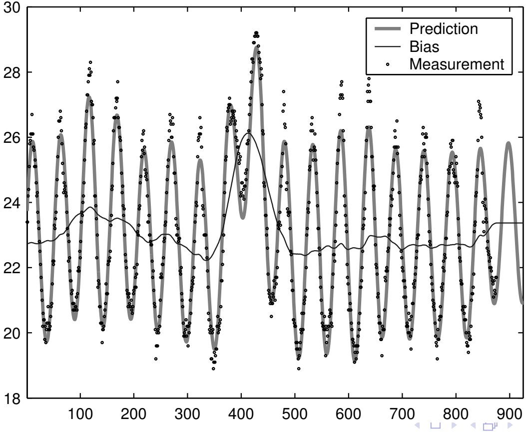
- Measurement model is of the form

$$y_k = \mathbf{H}\,\mathbf{x}_k + r_k$$

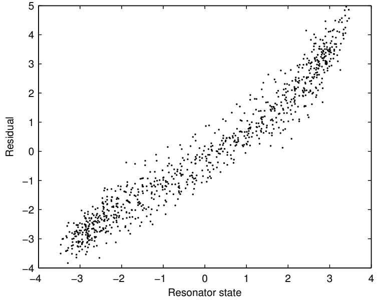
- Linear state space model $\Rightarrow$ **Kalman filter** can be applied.

Bias and Periodic Components 3/3

Finding the parameters that minimize the 50 step prediction error, results in the following kind of prediction:



Non-linear Correction Term

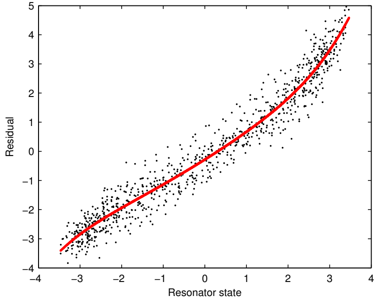


- Residual and periodic component still have a non-linear relationship
- 5th degree polynomial gives a suitable fit
- Measurement model becomes non-linear

$$y_k = x_b(r) + \sum_{i=0}^{5} c_i \left(x_r(k)\right)^i + r_k$$

- Use unscented Kalman filter (UKF) instead of Kalman filter

Non-linear Correction Term

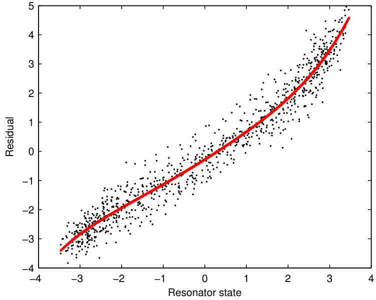


- Residual and periodic component still have a non-linear relationship
- 5th degree polynomial gives a suitable fit
- Measurement model becomes non-linear

$$y_k = x_b(r) + \sum_{i=0}^{5} c_i \left(x_r(k)\right)^i + r_k$$

- Use unscented Kalman filter (UKF) instead of Kalman filter

Non-linear Correction Term

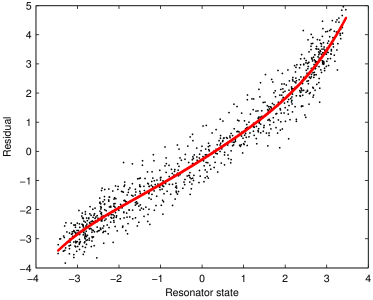


- Residual and periodic component still have a non-linear relationship
- 5th degree polynomial gives a suitable fit
- Measurement model becomes non-linear

$$y_k = x_b(r) + \sum_{i=0}^{5} c_i \left(x_r(k)\right)^i + r_k$$

- Use unscented Kalman filter (UKF) instead of Kalman filter

Non-linear Correction Term

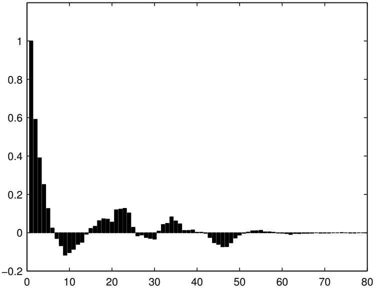


- Residual and periodic component still have a non-linear relationship
- 5th degree polynomial gives a suitable fit
- Measurement model becomes non-linear

$$y_k = x_b(r) + \sum_{i=0}^{5} c_i \left(x_r(k)\right)^i + r_k$$

- Use unscented Kalman filter (UKF) instead of Kalman filter

Autocorrelation Compensation

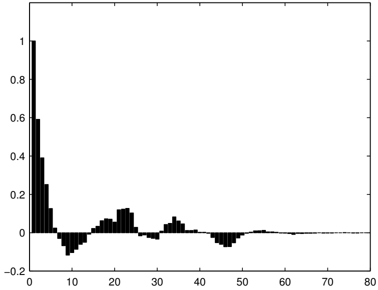


- Residual has non-delta autocorrelation, indicating a periodic component
- Can be modeled as second order AR-model:

$$e_k = \sum_{i=1}^{2} a_i \, e_{k-i} + r_k^{\mathrm{ar}}$$

- Can be estimated with Rauch-Tung-Striebel smoother

Autocorrelation Compensation

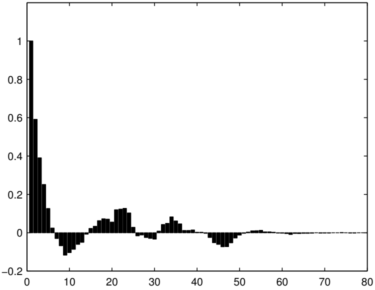


- Residual has non-delta autocorrelation, indicating a periodic component
- Can be modeled as second order AR-model:

$$e_k = \sum_{i=1}^{2} a_i \, e_{k-i} + r_k^{\mathtt{ar}}$$

- Can be estimated with Rauch-Tung-Striebel smoother

Autocorrelation Compensation



- Residual has non-delta autocorrelation, indicating a periodic component
- Can be modeled as second order AR-model:

$$e_k = \sum_{i=1}^{2} a_i \, e_{k-i} + r_k^{\mathrm{ar}}$$

- Can be estimated with Rauch-Tung-Striebel smoother

Estimation of parameters

- The unknown parameters are the spectral densities of process noises and the angular velocity of the resonator
- Form a discrete grid of sensible parameter values
- Evaluate parameters by computing 50 step prediction errors in known parts of time series
- First find roughly the location of minimum and form denser grid on that area
- Find the final smoothed estimate of the time series and make final prediction with the parameter values giving the minimum error

Estimation of parameters

- The unknown parameters are the spectral densities of process noises and the angular velocity of the resonator
- Form a discrete grid of sensible parameter values
- Evaluate parameters by computing 50 step prediction errors in known parts of time series
- First find roughly the location of minimum and form denser grid on that area
- Find the final smoothed estimate of the time series and make final prediction with the parameter values giving the minimum error

Estimation of parameters

- The unknown parameters are the spectral densities of process noises and the angular velocity of the resonator
- Form a discrete grid of sensible parameter values
- Evaluate parameters by computing 50 step prediction errors in known parts of time series
- First find roughly the location of minimum and form denser grid on that area
- Find the final smoothed estimate of the time series and make final prediction with the parameter values giving the minimum error

Estimation of parameters

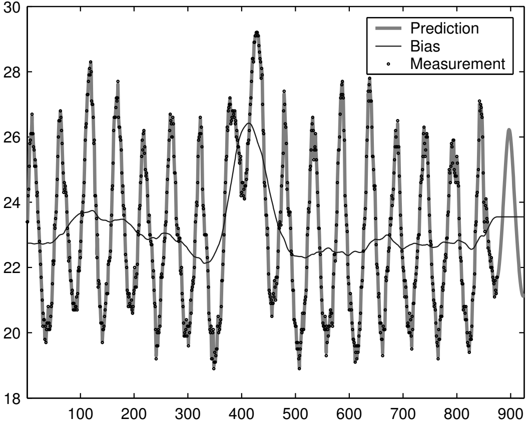
- The unknown parameters are the spectral densities of process noises and the angular velocity of the resonator
- Form a discrete grid of sensible parameter values
- Evaluate parameters by computing 50 step prediction errors in known parts of time series
- First find roughly the location of minimum and form denser grid on that area
- Find the final smoothed estimate of the time series and make final prediction with the parameter values giving the minimum error

Estimation of parameters

- The unknown parameters are the spectral densities of process noises and the angular velocity of the resonator
- Form a discrete grid of sensible parameter values
- Evaluate parameters by computing 50 step prediction errors in known parts of time series
- First find roughly the location of minimum and form denser grid on that area
- Find the final smoothed estimate of the time series and make final prediction with the parameter values giving the minimum error

Prediction

The final estimate of the signal and the prediction result:



Summary

- First the bias and periodic components are **modeled** as linear state space model
- Non-linear dependence between residual and periodic component is **modeled** with 5th degree polynomial
- Remaining autocorrelation is **modeled** with second order autoregressive (AR) model
- The **estimation** and **prediction** is done with unscented Kalman filter and Rauch-Tung-Striebel smoother
- Quite classical model based **(Bayesian)** approach, where the uncertainties are modeled as stochastic processes

Summary

- First the bias and periodic components are **modeled** as linear state space model
- Non-linear dependence between residual and periodic component is **modeled** with 5th degree polynomial
- Remaining autocorrelation is **modeled** with second order autoregressive (AR) model
- The **estimation** and **prediction** is done with unscented Kalman filter and Rauch-Tung-Striebel smoother
- Quite classical model based **(Bayesian)** approach, where the uncertainties are modeled as stochastic processes

Summary

- First the bias and periodic components are **modeled** as linear state space model
- Non-linear dependence between residual and periodic component is **modeled** with 5th degree polynomial
- Remaining autocorrelation is **modeled** with second order autoregressive (AR) model
- The **estimation** and **prediction** is done with unscented Kalman filter and Rauch-Tung-Striebel smoother
- Quite classical model based **(Bayesian)** approach, where the uncertainties are modeled as stochastic processes

Summary

- First the bias and periodic components are **modeled** as linear state space model
- Non-linear dependence between residual and periodic component is **modeled** with 5th degree polynomial
- Remaining autocorrelation is **modeled** with second order autoregressive (AR) model
- The **estimation** and **prediction** is done with unscented Kalman filter and Rauch-Tung-Striebel smoother
- Quite classical model based **(Bayesian)** approach, where the uncertainties are modeled as stochastic processes

Summary

- First the bias and periodic components are **modeled** as linear state space model
- Non-linear dependence between residual and periodic component is **modeled** with 5th degree polynomial
- Remaining autocorrelation is **modeled** with second order autoregressive (AR) model
- The **estimation** and **prediction** is done with unscented Kalman filter and Rauch-Tung-Striebel smoother
- Quite classical model based **(Bayesian)** approach, where the uncertainties are modeled as stochastic processes