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

1 Optimal Filtering and Smoothing
   - State Space Models
   - Optimal Filters and Predictors
   - Optimal Smoothers

2 Time Series Model
   - Idea of the Approach
   - Bias and Periodic Components
   - Non-linear Correction Term
   - Autocorrelation Compensation

3 Estimation of Parameters and Prediction Result
   - Estimation of Parameters
   - Final Prediction Result

4 Summary

Särkkä et al. Prediction of ESTSP Competition Time Series by UKF and ...
Contents

1 Optimal Filtering and Smoothing
   • State Space Models
   • Optimal Filters and Predictors
   • Optimal Smoothers

2 Time Series Model
   • Idea of the Approach
   • Bias and Periodic Components
   • Non-linear Correction Term
   • Autocorrelation Compensation

3 Estimation of Parameters and Prediction Result
   • Estimation of Parameters
   • Final Prediction Result

4 Summary

Särkkä et al. Prediction of ESTSP Competition Time Series by UKF and ...
1 Optimal Filtering and Smoothing
- State Space Models
- Optimal Filters and Predictors
- Optimal Smoothers

2 Time Series Model
- Idea of the Approach
- Bias and Periodic Components
- Non-linear Correction Term
- Autocorrelation Compensation

3 Estimation of Parameters and Prediction Result
- Estimation of Parameters
- Final Prediction Result

4 Summary

Särkkä et al. Prediction of ESTSP Competition Time Series by UKF and . . .
Optimal Filtering and Smoothing

1. State Space Models
2. Optimal Filters and Predictors
3. Optimal Smoothers

Time Series Model

2. Idea of the Approach
3. Bias and Periodic Components
4. Non-linear Correction Term
5. Autocorrelation Compensation

Estimation of Parameters and Prediction Result

3. Estimation of Parameters
4. Final Prediction Result

Summary
State Space Model

- **State space model** with state $x_k$ and measurements $y_k$:

$$x_k = f(x_{k-1}, q_{k-1})$$

$$y_k = h(x_k, r_k),$$

where $q_{k-1}$ is the process noise and $r_k$ is the measurement noise.

- The state $x_k$ is the hidden internal dynamic state of the system on the time step $k$.
- The measurements $y_k$ 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 $x_k$ and measurements $y_k$:

  \[
  x_k = f(x_{k-1}, q_{k-1}) \\
  y_k = h(x_k, r_k),
  \]

  where $q_{k-1}$ is the process noise and $r_k$ is the measurement noise.

- The state $x_k$ is the **hidden internal dynamic state** of the system on the time step $k$

- The measurements $y_k$ 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 \( x_k \) and measurements \( y_k \):

\[
\begin{align*}
  x_k & = f(x_{k-1}, q_{k-1}) \\
  y_k & = h(x_k, r_k),
\end{align*}
\]

where \( q_{k-1} \) is the process noise and \( r_k \) is the measurement noise.

- The state \( x_k \) is the **hidden internal dynamic state** of the system on the time step \( k \)

- The measurements \( y_k \) 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 with state $x_k$ and measurements $y_k$:

$$x_k = f(x_{k-1}, q_{k-1})$$
$$y_k = h(x_k, r_k),$$

where $q_{k-1}$ is the process noise and $r_k$ is the measurement noise.

- The state $x_k$ is the hidden internal dynamic state of the system on the time step $k$
- The measurements $y_k$ 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
Given measurements $y_1, \ldots, y_T$ optimal filter produces MMSE optimal online estimate:

$$\hat{x}(t_k) = E(x(t_k) | y_1, \ldots, y_k).$$

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

Can be also used for computing the optimal predictions:

$$\hat{x}(t) = E(x(t) | y_1, \ldots, y_k).$$

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 $y_1, \ldots, y_T$ optimal filter produces MMSE optimal online estimate:

  $$\hat{x}(t_k) = E(x(t_k) | y_1, \ldots, y_k).$$

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

- Can be also used for computing the optimal predictions:

  $$\hat{x}(t) = E(x(t) | y_1, \ldots, y_k).$$

  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 $y_1, \ldots, y_T$ optimal filter produces MMSE optimal online estimate:
  \[ \hat{x}(t_k) = E(x(t_k) | y_1, \ldots, y_k). \]
  for each $k = 1, \ldots, T$.
- Can be also used for computing the optimal predictions:
  \[ \hat{x}(t) = E(x(t) | y_1, \ldots, y_k). \]
  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.
Given measurements $y_1, \ldots, y_T$ optimal filter produces MMSE optimal online estimate:

$$\hat{x}(t_k) = E(x(t_k) \mid y_1, \ldots, y_k).$$

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

Can be also used for computing the optimal predictions:

$$\hat{x}(t) = E(x(t) \mid y_1, \ldots, y_k).$$

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:

  $$\hat{x}(t_k) = E(x(t_k) \mid y_1, \ldots, y_T).$$

- 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:

  \[ \hat{x}(t_k) = E(x(t_k) | y_1, \ldots, y_T). \]

- If the dynamic model is linear, then **Rauch-Tung-Striebel (RTS) smoother** provides the optimal solution.

- Approximate smoothers for nonlinear problems exist 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:
  \[
  \hat{x}(t_k) = E(x(t_k) | y_1, \ldots, y_T).
  \]

- 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:
  \[ \hat{x}(t_k) = E(x(t_k) | y_1, \ldots, y_T). \]

- 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.
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

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

\[
\frac{dx_b}{dt} = w_b(t)
\]

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

\[
\frac{d^2x_r}{dt^2} = -\omega^2 x_r + w_r(t)
\]

**Signal model:**

\[
y_k = x_b(k) + x_r(k) + r_k
\]
**Bias and Periodic Components**

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

\[
\frac{dx_b}{dt} = w_b(t)
\]

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

\[
\frac{d^2 x_r}{dt^2} = -\omega^2 x_r + w_r(t)
\]

**Signal model:**

\[
y_k = x_b(k) + x_r(k) + r_k
\]
Bias and Periodic Components

- **Bias** component $x_b(t)$ is modeled as integral of a white noise process $w_b(t)$
  \[
  \frac{dx_b}{dt} = w_b(t)
  \]

- **Periodic** component $x_r(t)$ is modeled as a white noise driven stochastic resonator
  \[
  \frac{d^2x_r}{dt^2} = -\omega^2 x_r + w_r(t)
  \]

**Signal model:**
\[
y_k = x_b(k) + x_r(k) + r_k
\]
Can be written as **discretely measured continuous-time vector process** $\mathbf{x}(t) = (\mathbf{x}_b(t) \ \mathbf{x}_r(t) \ \mathrm{d}\mathbf{x}_r(t)/\mathrm{d}t)^T$ 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 \( x(t) = (x_b(t) \ x_r(t) \ dx_r(t)/dt)^T \) as follows:

\[
\frac{dx}{dt} = F \ x(t) + L \ w(t)
\]

- Linear system theory \( \Rightarrow \) equivalent discrete time model:

\[
x_k = A \ x_{k-1} + q_{k-1}
\]

- Measurement model is of the form

\[
y_k = H \ x_k + r_k
\]

- Linear state space model \( \Rightarrow \) Kalman filter can be applied.
Can be written as \textit{discretely measured continuous-time vector process} \( x(t) = (x_b(t) \ x_r(t) \ dx_r(t)/dt)^T \) as follows:

\[
\frac{dx}{dt} = F x(t) + L w(t)
\]

Linear system theory \( \Rightarrow \) equivalent \textit{discrete time model}:

\[
x_k = A x_{k-1} + q_{k-1}
\]

Measurement model is of the form

\[
y_k = H x_k + r_k
\]

Linear state space model \( \Rightarrow \) \textit{Kalman filter} can be applied.
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:

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

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

$$y_k = \mathbf{H} \mathbf{x}_k + \mathbf{r}_k$$

Linear state space model $\Rightarrow$ Kalman filter can be applied.
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
Residual has **non-delta autocorrelation**, indicating a periodic component.

Can be modeled as **second order AR-model**:

\[ e_k = \sum_{i=1}^{2} a_i e_{k-i} + r_{k}^{ar} \]

Can be estimated with **Rauch-Tung-Striebel smoother**.
Residual has **non-delta autocorrelation**, indicating a periodic component.

Can be modeled as a **second order AR-model**:

\[ e_k = \sum_{i=1}^{2} a_i e_{k-i} + r_{ar}^k \]

Can be estimated with the **Rauch-Tung-Striebel smoother**.
Residual has non-delta autocorrelation, indicating a periodic component.
Can be modeled as second order AR-model:
\[
e_k = \sum_{i=1}^{2} a_i e_{k-i} + r_k^{ar}
\]
Can be estimated with Rauch-Tung-Striebel smoother.
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.
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.
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.
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.
The final estimate of the signal and the prediction result:
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 a 5th degree polynomial.
- Remaining autocorrelation is modeled with a 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.

Prediction of ESTSP Competition Time Series by UKF and . . .
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.