Functional estimation in systems defined by differential equations using bayesian smoothing methods

January 30th 2008

Model defined by differential equations

Definition

Definition

$$\begin{cases} Dx(t) & = & f(x(t), u(t), \theta) \\ x(0) & = & x_0 \end{cases}$$

With:

- $x(t)$ the set of d **output** functions,
- $u(t)$ the set of q **input** functions,
- θ the **vector of parameter** involved the set of differential equations.

Existence & uniqueness of the solution

f **Lipchitz continuous** and $u(t)$ **differentiable** almost everywhere

Observations

Output functions observed at time points t_i for $i = 1, \ldots, n$ with **measurement errors** ε_i:

$$y_i = x(t_i) + \varepsilon_i \quad i = 1, \ldots, n$$

Area & Current methods

Area of application

- Chemical engineering,
- Pharmacokinetic / pharmacodynamic,
- …

Objectives

- Estimate the **vector of parameters** θ,
- Estimate the **output function** $x(t)$.

Data fitting by numerical approximation of an initial value problem

- **Approximation** of the output function using **numerical methods** (e.g. Runge-Kutta algorithm),
- **Updates** of parameter estimate using this fitted curve into an **optimization algorithm.**

Limitations & problems

- Computationally very intensive,
- Problem of **instability**.

Bayesian Smoothing methods for differential equation models

B-splines definition & properties

Definition

B-spline basis function defined using:

- order p,
- m inner knots at $\tau_1 \leq \cdots \leq \tau_m$,
- p-multiple knots τ_0 and τ_{m+1},
- **Recursive definition** for each function $B_k(t,p)$.

Properties

- $B_k(t,p)$ is a **piecewise polynomial** of degree $p-1$,
- Derivatives up to order $p-2$ are continuous,
- **Sum** of all non-zero basis function is **1**,
- Number of basis function is $\boldsymbol{K = m + p}$.

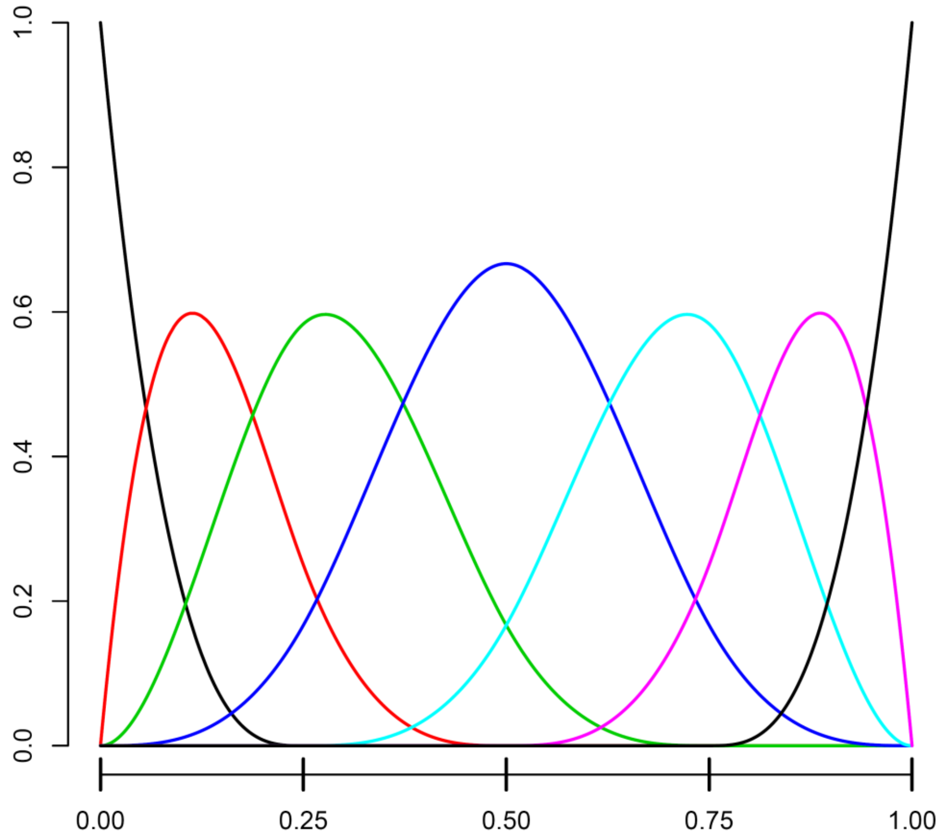

Figure 1 : B-spline basis of order 4 with 3 inner knots

Generalized Profiling for DE

Basis function expansion

$$\hat{x}(t) = \sum_{k=1}^{K} c_k B_k(t,p) = c^T B(t)$$

Fitting criterion

$$J(c,\theta|\lambda,y) = \sum_{i=1}^{n} \left(y_i - \hat{x}(t_i)\right)^2 + \lambda \int \left(L_\theta \hat{x}(t)\right)^2 dt$$
$$= \|y - Bc\|^2 + \lambda * PEN(\hat{x})$$

With:

- $L_\theta\left(x(t)\right) = Dx(t) - f(x(t), u(t), \theta)$ the **differential equation operator**,
- λ corresponds to **weight fidelity term**.

Linear case

$$PEN(\hat{x}) = c^T R(\theta) c$$

With $R(\theta) = \int L_\theta B(t) * \left(L_\theta B(t)\right)^T dt$

Bayesian Generalized Profiling for DE

Bayesian model

$$\begin{cases} y(t)|c,\tau & \sim & \mathcal{N}(c^T B(t);\tau) \\ \pi(c|\gamma,\theta) & \propto & \exp\left(-\frac{\gamma}{2} PEN(\hat{x})\right) \\ \gamma & \sim & \mathcal{G}a(a_\gamma; b_\gamma) \\ \tau & \sim & \mathcal{G}a(a_\tau; b_\tau) \\ \theta & \sim & \pi(\theta) \end{cases}$$

Joint log-posterior in the linear case

$$\begin{aligned} l(c,\gamma,\tau,\theta|I,y) \triangleq\ & \frac{n}{2}\log(\tau) - \frac{\tau}{2}(y - Bc)^T(y - Bc) + \\ & \frac{\rho\big(R(\theta)\big)}{2}\log(\gamma) + \frac{1}{2}\log(|R(\theta)|) - \frac{\gamma}{2} c^T R(\theta) c + \\ & \left(a_\gamma - 1\right)\log(\gamma) - b_\gamma \gamma + \\ & (a_\tau - 1)\log(\tau) - b_\tau \tau + \\ & \log\big(\pi(\theta)\big) \end{aligned}$$

Bayesian Generalized Profiling for DE

Conditional posteriors

$$\begin{cases} \tau|c,y & \sim & \mathcal{G}a\left(\frac{n}{2}+a_\tau; \frac{(y-Bc)^T(y-Bc)}{2}+b_\tau\right) \\ \gamma|\theta,c,y & \sim & \mathcal{G}a\left(\frac{\rho\left(R(\theta)\right)}{2}+a_\gamma; \frac{c^T R(\theta)c}{2}+b_\gamma\right) \\ c|\theta,\gamma,c,y & \sim & \mathcal{N}\left(\left(B^TB+\frac{\gamma}{\tau}R(\theta)\right)^{-1}B^Ty; \left(\tau B^TB+\gamma R(\theta)\right)^{-1}\right) \\ \pi(\theta|c,\gamma,y) & \propto & |R(\theta)|^{\frac{1}{2}} * \exp\left(-\frac{\gamma}{2}c^TR(\theta)c\right) * \pi(\theta) \end{cases}$$

Possibility for **Gibbs sampling** for the parameter τ, γ and c
Unknown distribution for the conditional posterior of θ

Solution

Metropolis-Hastings within Gibbs algorithm

Illustration

I. Model defined by differential equations

II. Bayesian Smoothing methods for differential equation models

III. **Illustration**

IV. Further work & conclusion

Exponential Decline

Simple linear differential equation

$$\begin{cases} Dx(t) & = & -\theta x(t) \quad t \in [0; T] \\ x(0) & = & 1 \end{cases}$$

Analytic solution

$$x(t) = \exp(-\theta t)$$

Generating data

$n = 50$ measurments
θ positiv, e.g. $\theta = 5$
Additive Gaussian error with $\sigma = 0.1$

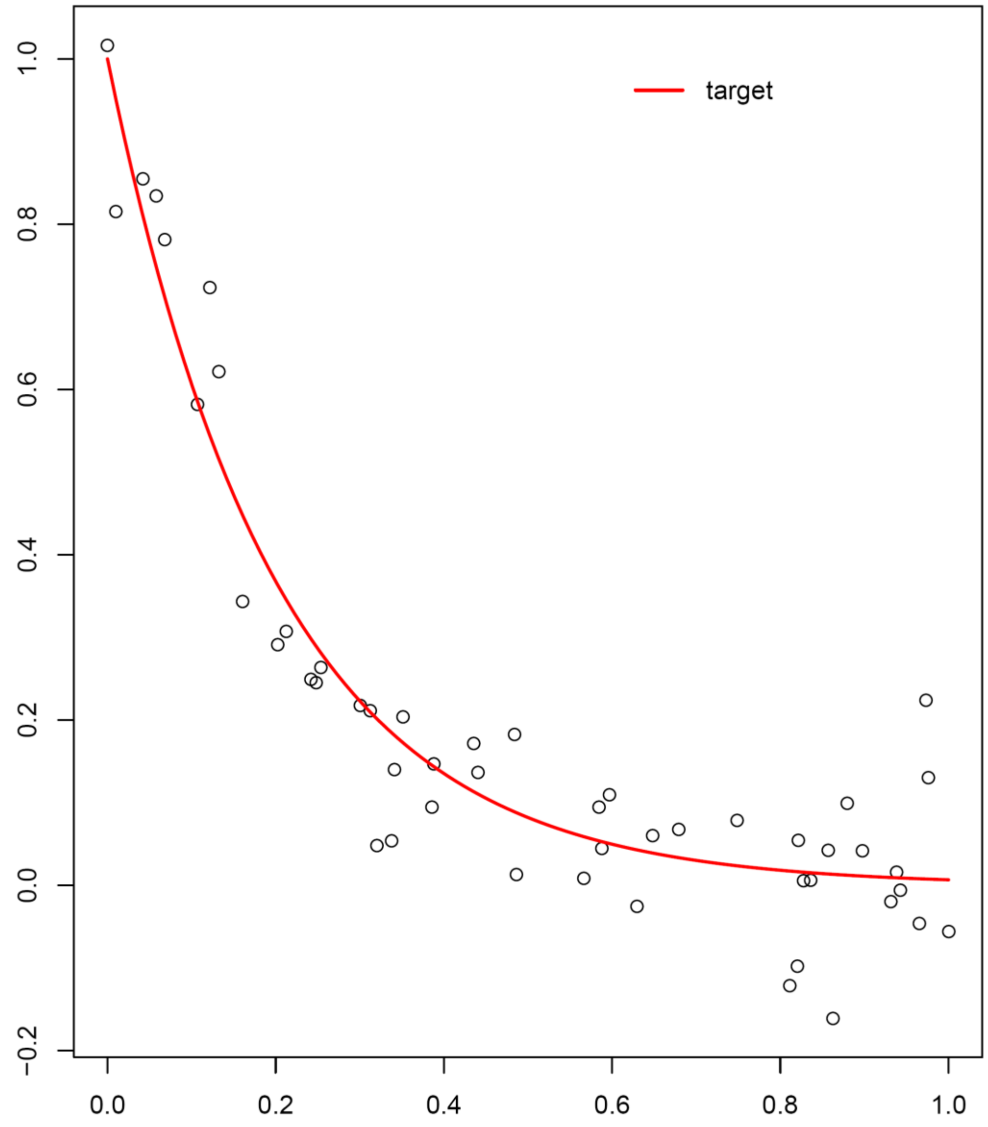

Figure 2 - Simulated data & true curve

Exponential Decline

Simple linear differential equation

$$\begin{cases} Dx(t) & = & -\theta x(t) \quad t \in [0;T] \\ x(0) & = & 1 \end{cases}$$

Analytic solution

$$x(t) = \exp(-\theta t)$$

Generating data

$n = 50$ measurments
θ positiv, e.g. $\theta = 5$
Additive Gaussian error with $\sigma = 0.1$

Result with the generalized profiling for DE

$\hat{\theta} = 5.186$
$\hat{\sigma} = 8.058E - 2$

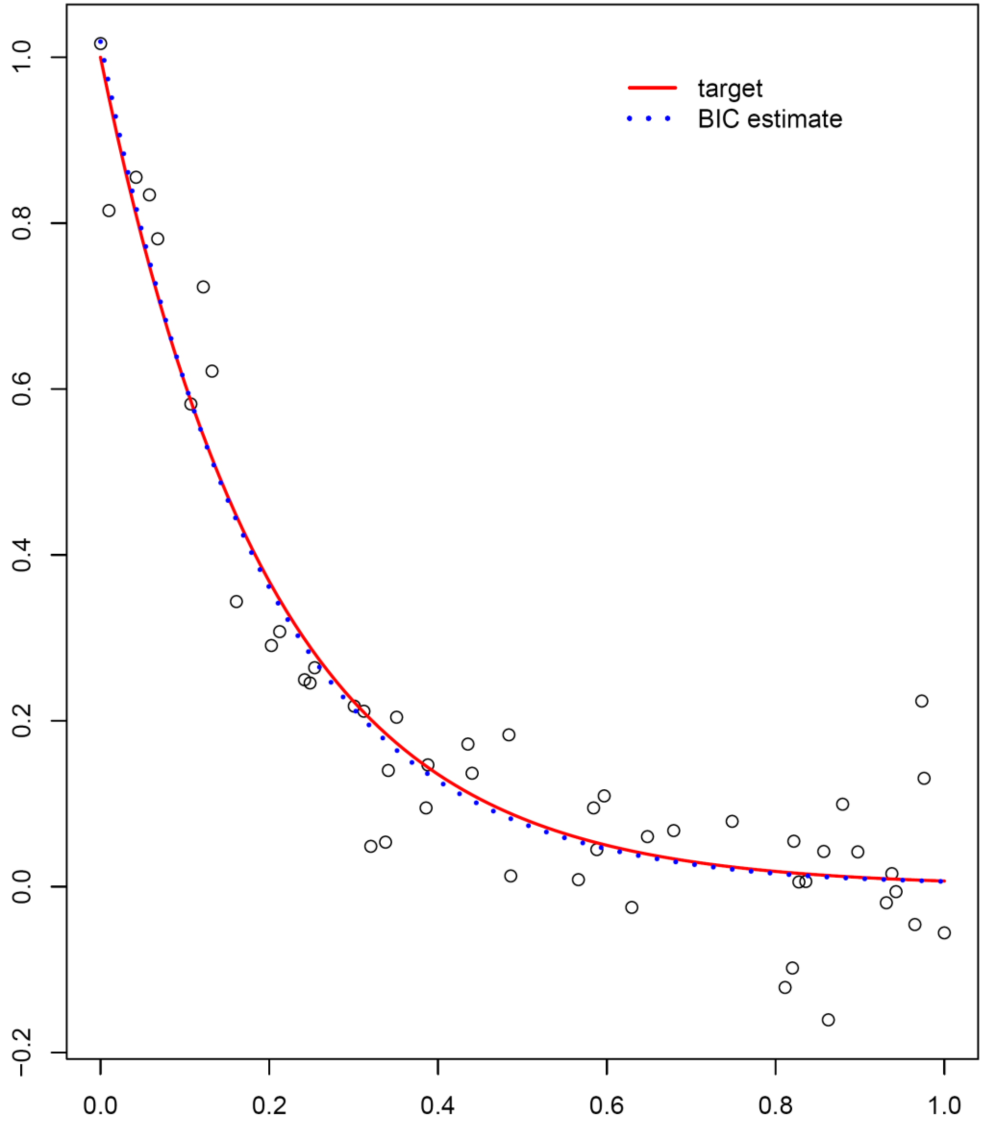

Figure 3 - Simulated data, true curve & estimate curve using BIC criterion

Exponential Decline

Bayesian Approach

$$\begin{cases} y(t)|c,\tau & \sim & \mathcal{N}(c^T B(t);\tau) \\ \pi(c|\gamma,\theta) & \propto & \exp\left(-\frac{\gamma}{2}c^T R(\theta)c\right) \\ \gamma & \sim & \mathcal{G}a(a_\gamma; b_\gamma) \\ \tau & \sim & \mathcal{G}a(a_\tau; b_\tau) \\ \theta & \sim & \mathcal{LN}(\mu_\theta, \tau_\theta) \end{cases}$$

Convergence diagnostics

Convergence of all the chains
High **autocorrelations** in the θ chain.
Slow mixing and **slow convergence** for θ.

Estimation

$\hat{\theta} = 5.287$ with $var(\hat{\theta}) = 0.122$
$\sigma = 8.27E-2$

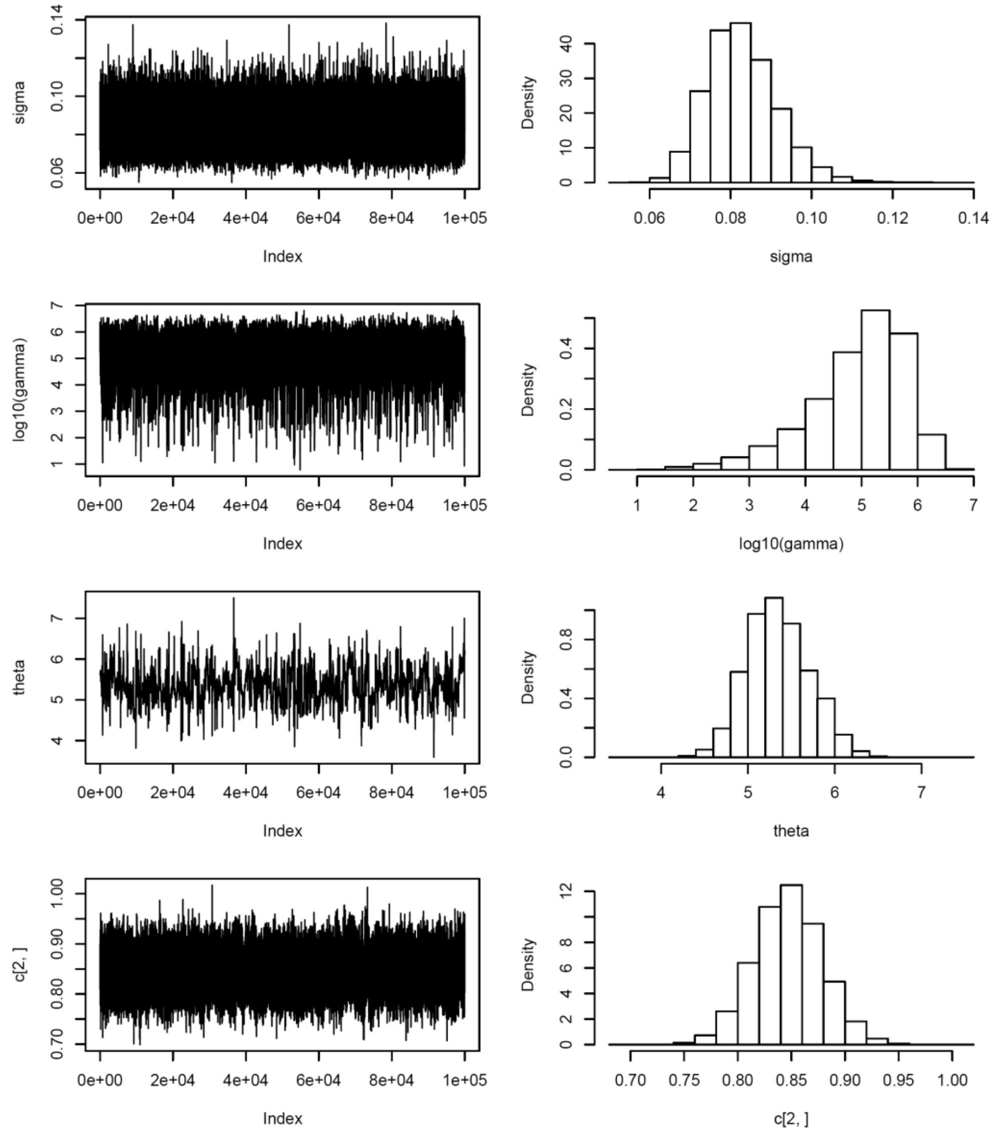

Figure 4 – Trace for the parameters σ, $\log_{10}(\gamma)$, θ and c_2.

Exponential Decline

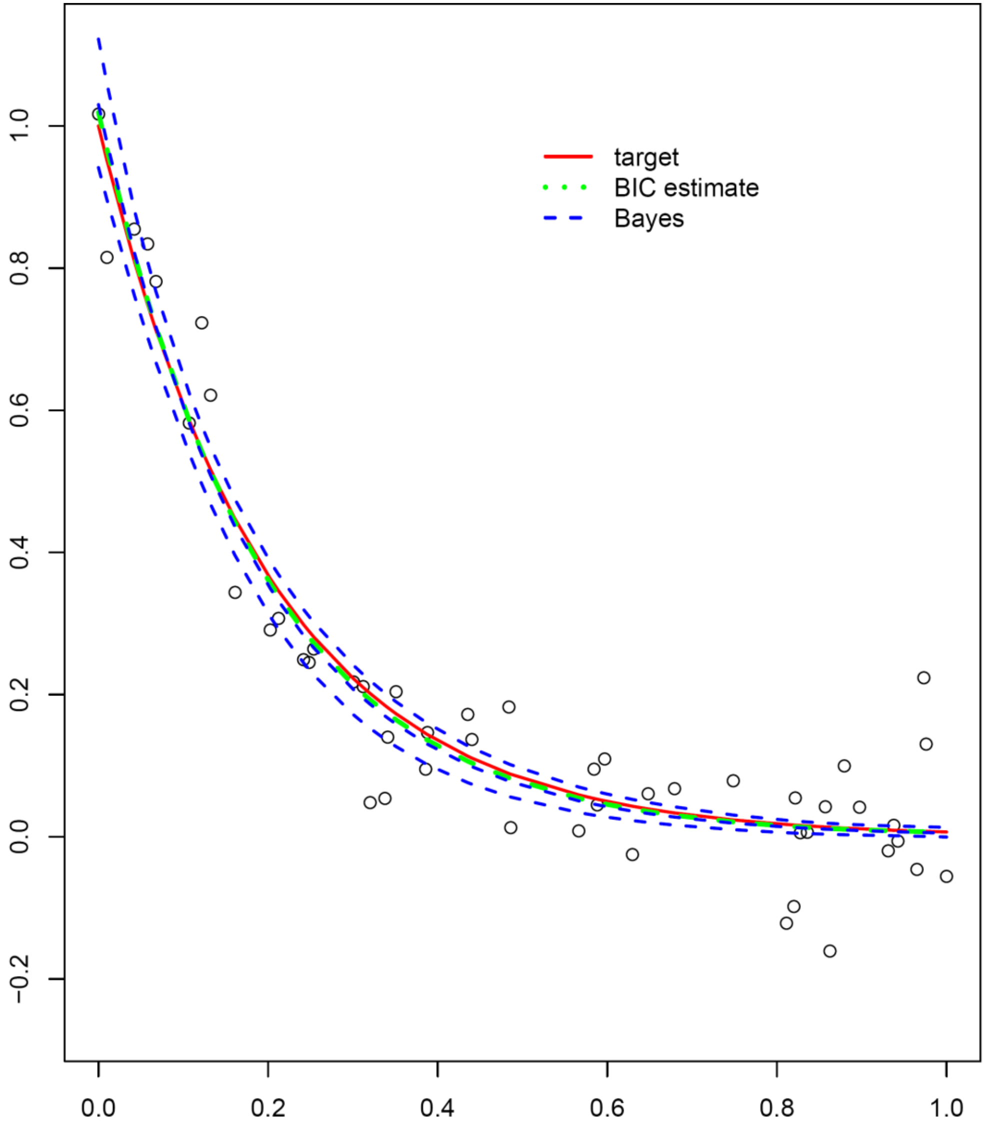

Figure 5 - 95% posterior credibility interval for the conditional posterior mean

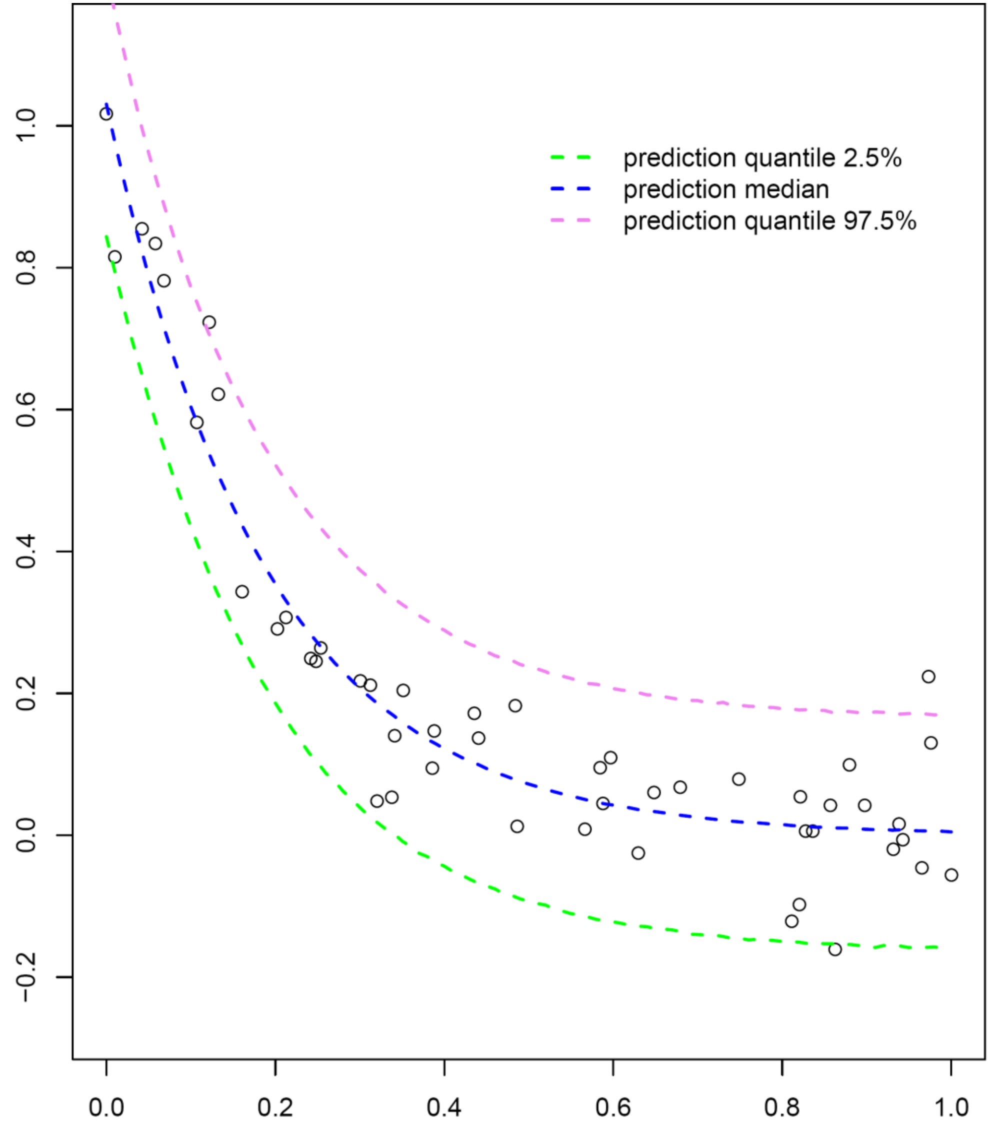

Figure 6 - Visual Predictive Check

Further work & conclusion

I. Model defined by differential equations

II. Bayesian Smoothing methods for differential equation models

III. Illustration

IV. Further work & conclusion

Further work & conclusion

Conclusion

- **Powerful tool** which overcomes the numerical integrations in the current method,
- **Convenient implementation** of the Bayesian generalized profiling for DE,
- Simple method to include **prior information**,
- Possibility to express **uncertainty** with respect to initial conditions.

Further work

- Propose a method for **mixed effect model**,
- Generalize this method to **nonlinear differential equations**,
- **Optimal design** for the data collection,
- Differential equation model with **lagged effects** e.g. $Dx(t) = f(x(t-\delta_1), u(t-\delta_2), \theta)$,
- **Stochastic differential equation** with Brownian motion.

References

[1] Berry S.M., Carroll R.J. and Ruppert D., Bayesian smoothing and regression splines for measurement error problems, Journal of the American Statistical Association, **97**:160-169 (2002)

[2] Poyton A.A, Varziri M.S., McAuley K.B., McLellan P.J. and Ramsay J.O., Parameter estimation in continuous-time dynamic models using principal differential analysis, Computers and Chemical Engineering, **30**:698-708 (2006)

[3] Ramsay J.O., Hooker G., Campbell D. and Cao J., Parameter estimation for differential equations: a generalized smoothing approach, Journal of the Royal Statistical Society, Series B, **69**:741-796 (2007)

[4] Campbell D., Bayesian collocation tempering and generalized profiling for estimation of parameters from differential equation models, PhD Thesis (2007)