



**HAL**  
open science

# AUTOREGRESSIVE MODEL WITH PARTIAL FORGETTING WITHIN RAO-BLACKWELLIZED PARTICLE FILTER

Kamil Dedecius, Radek Hofman

► **To cite this version:**

Kamil Dedecius, Radek Hofman. AUTOREGRESSIVE MODEL WITH PARTIAL FORGETTING WITHIN RAO-BLACKWELLIZED PARTICLE FILTER. *Communications in Statistics - Simulation and Computation*, 2011, 41 (05), pp.582-589. 10.1080/03610918.2011.598992 . hal-00768970

**HAL Id: hal-00768970**

**<https://hal.science/hal-00768970>**

Submitted on 27 Dec 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**AUTOREGRESSIVE MODEL WITH PARTIAL FORGETTING  
WITHIN RAO-BLACKWELLIZED PARTICLE FILTER**

Journal:	<i>Communications in Statistics - Simulation and Computation</i>
Manuscript ID:	LSSP-2011-0055
Manuscript Type:	Original Paper
Date Submitted by the Author:	09-Feb-2011
Complete List of Authors:	Dedecius, Kamil; Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Department of Adaptive Systems Hofman, Radek; Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Department of Adaptive Systems
Keywords:	particle filters, Bayesian methods, recursive estimation
Abstract:	We are concerned with Bayesian identification and prediction of a nonlinear discrete stochastic process. The fact, that a nonlinear process can be approximated by a piecewise linear function advocates the use of adaptive linear models. We propose a linear regression model within Rao-Blackwellized particle filter. The parameters of the linear model are adaptively estimated using a finite mixture, where the weights of components are tuned with a particle filter. The mixture reflects a priori given hypotheses on different scenarios of (expected) parameters' evolution. The resulting hybrid filter locally optimizes the weights to achieve the best fit of signal.
Note: The following files were submitted by the author for peer review, but cannot be converted	

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

comm.stat.tex

SCHOLARONE™  
Manuscripts

For Peer Review Only

1 AUTOREGRESSIVE MODEL WITH PARTIAL FORGETTING WITHIN RAO-BLACKWELLIZED PARTICLE  
2 FILTER  
3  
4  
5

6 Kamil Dedecius<sup>1,2</sup> and Radek Hofman<sup>1</sup>

7  
8 <sup>1</sup> Department of Adaptive Systems

9 Institute of Information Theory and Automation

10 Academy of Sciences of the Czech Republic

11 Pod Vodarenskou vezi 4, 182 08 Prague, Czech Republic

12  
13  
14 <sup>2</sup> Department of Applied Mathematics

15 Faculty of Transportation Sciences

16 Czech Technical University in Prague

17 Na Florenci 25, 110 00 Prague, Czech Republic

18  
19  
20  
21 {dedecius,hofman}@utia.cas.cz

22  
23  
24 Key Words: particle filters; Bayesian methods; recursive estimation

25  
26 ABSTRACT

27 We are concerned with Bayesian identification and prediction of a nonlinear discrete stochastic process. The fact,  
28 that a nonlinear process can be approximated by a piecewise linear function advocates the use of adaptive linear  
29 models. We propose a linear regression model within Rao-Blackwellized particle filter. The parameters of the linear  
30 model are adaptively estimated using a finite mixture, where the weights of components are tuned with a particle  
31 filter. The mixture reflects *a priori* given hypotheses on different scenarios of (expected) parameters' evolution. The  
32 resulting hybrid filter locally optimizes the weights to achieve the best fit of a nonlinear signal with a single linear  
33 model.  
34  
35  
36  
37

38  
39 1. INTRODUCTION

40 The theory of approximation of nonlinear signals by piecewise linear functions has attained considerable attention  
41 over the recent decades (Tong, 1993), mainly in the field of control engineering, e.g., Kalman (1957). The method  
42 is popular, because in many technical applications, it offers a reasonable trade-off between the model complexity and  
43 performance (Sontag 2002; Savona 1988). Our method is motivated by the fact that if the transitions between two  
44 successive (almost) linear segments is smooth enough, it can be modelled with switching models, e.g., Judd and Mees  
45 (1995), Ragot *et al.* (2004), Rosenqvist and Karlström (2005), or model averaging, Raftery *et al.* (2010).  
46  
47  
48  
49

50 Notational conventions:  $\propto$  denotes proportionality, i.e., equality up to a constant factor.  $A'$  denotes transpose of  $A$ .  
51  $p(a|b)$  is a probability density function (pdf) of a (multivariate) random variable  $a$  given  $b$ . The pdfs are distinguished  
52  
53  
54  
55

by their argument.  $t \in \{1, 2, \dots\}$  denotes discrete time instants. All integrations are over the maximal plausible support.

## 2. BAYESIAN APPROACH TO MODELLING

Assume that we are given a time series of real observations  $\mathbf{Y}^{t-1} = (y_1, \dots, y_{t-1})$  and our purpose is to determine its next value  $y_t$ . The statistical approach employs, among others, the parametric models describing the dependence of  $y_t$  on previous observation  $\mathbf{Y}^{t-1}$  through conditional distributions pdf

$$p(y_t | \mathbf{Y}^{t-1}, \Theta). \quad (1)$$

Under the Bayesian treatment,  $\Theta$  is a set of constant model parameters with pdf  $p(\Theta | \mathbf{Y}^{t-1})$ . If this distribution is properly chosen from a class conjugate to the model (1), the Bayes' theorem yields a posterior pdf of the same type, and the recursive data update reads (Bernardo and Smith, 2001; Peterka, 1981)

$$p(\Theta | \mathbf{Y}^t) = \frac{p(y_t | \mathbf{Y}^{t-1}, \Theta) p(\Theta | \mathbf{Y}^{t-1})}{p(y_t | \mathbf{Y}^{t-1})}. \quad (2)$$

The predictive pdf  $p(y_{t+1} | \mathbf{Y}^t)$  provides the Bayesian output prediction. Using the Chapman-Kolmogorov equation (Karush 1961), it holds that

$$p(y_{t+1} | \mathbf{Y}^t) = \int p(y_{t+1} | \mathbf{Y}^t, \Theta) p(\Theta | \mathbf{Y}^t) d\Theta = \frac{I_{t+1}}{I_t}, \quad (3)$$

where  $I$  denotes the normalizing integral, see, e.g., Peterka (1981).

Although the described methodology is important *per se*, its lack of adaptivity prevents successful application to nonstatic cases, when  $\Theta$  is not time-invariant. For a time varying  $\Theta_t$ , it is necessary to perform an additional update

$$p(\Theta_t | \mathbf{Y}^t) \rightarrow p(\Theta_{t+1} | \mathbf{Y}^t). \quad (4)$$

Here, two significant cases can occur:

1. The evolution model  $p(\Theta_{t+1} | \Theta_t, \mathbf{Y}^t)$  is known *a priori*.
2. A suitable model of parameter evolution is not known, but we can expect that they vary slowly.

Case (i) allows one to aggregate both (1) and  $p(\Theta_{t+1} | \Theta_t, \mathbf{Y}^t)$  into a single complex model. Then,  $\Theta_t$  represents the system state and, under certain conditions, the estimation task leads to the well known Kalman filter, Peterka (1981).

However, there exists a wide variety of cases, when the explicit model of parameter evolution is not known. If we adopt the assumption of slowly varying parameters (ii), the popular group of estimation methods using forgetting provides a solution. It heuristically circumvents the problem of parameter model ignorance by recursive discounting the old and potentially outdated information carried by the parameter pdf. Formally, we introduce a forgetting operator  $\tilde{\mathfrak{F}}$ :

$$p(\Theta_{t+1} | \mathbf{Y}^t) = \tilde{\mathfrak{F}}[p(\Theta_t | \mathbf{Y}^t)]. \quad (5)$$

The application of the forgetting operator is equivalent to the time update in state-space models, Kárný (2006).

### 3. PARTIAL FORGETTING

The use of “classical” forgetting methods, e.g., exponential forgetting (Peterka, 1981), is limited in nonlinear cases. We present a new approach appealing to partial forgetting, Dedecius (2010). While it enumerates hypotheses about variability of elements of  $\Theta_t$ , our modification is more general.

Let us define a finite set  $\mathcal{H}$  of hypotheses  $\{H_i\}$  regarding the distribution of  $\Theta_{t+1}$  given  $\Theta_t$ . The distributions induced by these hypotheses are merged together in form of a finite mixture,

$$p(\Theta_{t+1}|\mathbf{Y}^t, \boldsymbol{\pi}_t) = \sum_i \pi_{i,t} q_i(\Theta_{t+1}|\mathbf{Y}^t), \quad \sum_i \pi_{i,t} = 1, \quad (6)$$

i.e., the posterior distribution of  $\Theta_{t+1}$  is represented by a finite mixture of hypothetical pdfs  $q_i(\Theta_{t+1}|\mathbf{Y}^t)$ .

A theoretically correct solution would express one hypothetical pdf for each (almost) linear window. In practice, this is hardly possible and a generalization of the approach is exploited. The hypotheses  $H_i \in \mathcal{H}$  enumerate several cases, that are likely to occur. Their number depends on a specific task, mainly the signal properties, and the user’s ability to guess the properties of  $\Theta_t$  in each window. For instance, we may state a hypothesis about each particular element of  $\Theta_t = \{\Theta_{1,t}, \dots, \Theta_{N,t}\}$  and about all of them in one shot:

$$\begin{aligned} H_0: \quad & p(\Theta_{t+1}|\mathbf{Y}^t, H_0) = p(\Theta_t|\mathbf{Y}^t) \\ & = q_0(\Theta_{t+1}|\mathbf{Y}^t) \\ H_1: \quad & p(\Theta_{t+1}|\mathbf{Y}^t, H_1) = p(\Theta_{2,t}, \dots, \Theta_{N,t}|\mathbf{Y}^t, \Theta_{1,t}) \\ & \quad \times \mathfrak{F}[p(\Theta_{1,t}|\mathbf{Y}^t)] = q_1(\Theta_{t+1}|\mathbf{Y}^t) \\ & \quad \vdots \\ H_N: \quad & p(\Theta_{t+1}|\mathbf{Y}^t, H_N) = p(\Theta_{1,t}, \dots, \Theta_{N-1,t}|\mathbf{Y}^t, \Theta_{N,t}) \\ & \quad \times \mathfrak{F}[p(\Theta_{N,t})] = q_N(\Theta_{t+1}|\mathbf{Y}^t) \\ H_{N+1}: \quad & p(\Theta_{t+1}|\mathbf{Y}^t, H_{N+1}) = \mathfrak{F}[p(\Theta_t|\mathbf{Y}^t)] \\ & = q_{N+1}(\Theta_{t+1}|\mathbf{Y}^t) \end{aligned}$$

This particular set of  $N + 1$  hypotheses is an example of many possible choices. The two extreme hypotheses  $H_0$  and  $H_{N+1}$  represent the user’s belief that none or all parameter vary, respectively. The remaining hypotheses  $H_1, \dots, H_N$  concern the case of variability of  $\Theta_t$ ’s one element (with appropriate index). We can choose different operators  $\mathfrak{F}$  or completely expert pdfs  $q_i(\Theta_{t+1}|\mathbf{Y}^t)$  as well.

Working with the mixture (6) would require a rather complex treatment. Instead, we prefer to find a single pdf  $\tilde{p}$  of the same class as the components, minimizing the expected Kullback-Leibler divergence of  $p$  on  $\tilde{p}$ , given by

$$\begin{aligned} \mathbb{E} [\mathcal{D}(p||\tilde{p}) | \mathcal{H}, \boldsymbol{\pi}_t, \mathbf{Y}^t] &= \\ &= \mathbb{E} \left[ \underbrace{\int p(\Theta_{t+1}|\mathbf{Y}^t, \boldsymbol{\pi}_t) \frac{p(\Theta_{t+1}|\mathbf{Y}^t, \boldsymbol{\pi}_t)}{\tilde{p}(\Theta_{t+1}|\mathbf{Y}^t, \boldsymbol{\pi}_t)} d\Theta}_{\rightarrow \min} | \mathcal{H}, \boldsymbol{\pi}_t, \mathbf{Y}^t \right]. \end{aligned} \quad (7)$$

It can be shown, that  $\mathcal{D}(f||g) \geq 0$  with equality for  $f = g$  almost everywhere, Kullback and Leibler (1951). A solution, for certain analytical cases (Dedecius, 2010), defines the approximate pdf  $\tilde{p}(\Theta_{t+1}|\mathbf{Y}^t, \pi_t)$ , which may be directly used as the next prior distribution pdf in (2).

#### 4. RAO-BLACKWELLIZED PARTICLE FILTER

Let  $\Psi_{t+1} = (\Theta'_{t+1}, \pi'_t)'$  be a real column vector. Given the value of  $\pi_t$ , the minimization (7) yielding the approximation  $\tilde{p}(\Theta_{t+1}|\mathbf{Y}^t, \pi_t)$  of (6), can be evaluated. The approximate pdf can be used for linear recursive estimation of model (1). Since weights  $\pi_t$  are unknown, we attempt to estimate the joint pdf of regression parameters and weights  $p(\Psi_{t+1}|\mathbf{Y}^t)$ . We exploit the natural factorization of  $\Psi_{t+1}$  and decompose the pdf  $p(\Psi_{t+1}|\mathbf{Y}^t)$  as follows

$$p(\Psi_{t+1}|\mathbf{Y}^t) = \underbrace{p(\Theta_{t+1}|\mathbf{Y}^t, \pi_t)}_{\text{linear}} \underbrace{p(\pi_t|\mathbf{Y}^t)}_{\text{PF}} \quad (8)$$

where  $p(\Theta_{t+1}|\mathbf{Y}^t, \pi_t)$  is analytically tractable while  $p(\pi_t|\mathbf{Y}^t)$  is not. The latter pdf is approximated using particle filter (PF) theory, Doucet *et al.* (2001).

Particle filtering refers to a range of techniques for generating an empirical approximation of the pdf

$$p(\Pi^t|\mathbf{Y}^t) \approx \frac{1}{M} \sum_{j=1}^M \delta(\Pi^t - \Pi^{t,(j)}), \quad (9)$$

where  $\Pi^t = (\pi_1, \dots, \pi_t)$  and  $\Pi^{t,(j)}$ ,  $j = 1, \dots, M$  are independent identically distributed samples from the posterior, and  $\delta(\cdot)$  denotes the Dirac  $\delta$ -function. Therefore, this approach is feasible only if we can sample from the exact posterior  $p(\Pi^t|\mathbf{Y}^t)$ . If this is not the case, the samples can be drawn from a chosen proposal distribution (importance function),  $f(\Pi^t|\mathbf{Y}^t)$ , as follows:

$$\begin{aligned} p(\Pi^t|\mathbf{Y}^t) &= \frac{p(\Pi^t|\mathbf{Y}^t)}{f(\Pi^t|\mathbf{Y}^t)} f(\Pi^t|\mathbf{Y}^t) \\ &\approx \frac{p(\Pi^t|\mathbf{Y}^t)}{f(\Pi^t|\mathbf{Y}^t)} \frac{1}{M} \sum_{j=1}^M \delta(\Pi^t - \Pi^{t,(j)}). \end{aligned} \quad (10)$$

Using the properties of the Dirac  $\delta$ -function, the approximation can be written in the form of a weighted empirical distribution, as follows:

$$p(\Pi^t|\mathbf{Y}^t) \approx \sum_{j=1}^M w_t^{(j)} \delta(\Pi^t - \Pi^{t,(j)}), \quad (11)$$

$$w_t^{(j)} \propto \frac{p(\Pi^{t,(j)}|\mathbf{Y}^t)}{f(\Pi^{t,(j)}|\mathbf{Y}^t)}. \quad (12)$$

Under this importance sampling procedure, the true posterior distribution needs only to be evaluated pointwise.

The challenge for on-line algorithms is to achieve recursive generation of samples and evaluation of the importance weights. Using standard Bayesian calculus, (12) may be written in the following recursive form:

$$w_t^{(j)} \propto \frac{p(y_{t+1}|\mathbf{Y}^t) p(\pi_t^{(j)}|\pi_{t-1}^{(j)})}{f(\pi_t^{(j)}|\Pi^{t-1,(j)}, \mathbf{Y}^t)} w_{t-1}^{(j)} \quad (13)$$

Furthermore, if  $f(\pi_t^{(j)} | \mathbf{Y}^{t-1}, \mathbf{Y}^t) = p(\pi_t^{(j)} | \pi_{t-1}^{(j)})$ , then the importance density becomes only dependent on the  $\pi_{t-1}$  and  $\mathbf{y}_t$ . This is particularly useful in the common case when only a filtered estimate of the posterior  $p(\pi_t | \mathbf{Y}^t)$  is required at each time step. It means that only the  $\pi_t^{(j)}$  need to be stored, Ristic *et al* (2004). Then, the marginal posterior density  $p(\pi_t | \mathbf{Y}^t)$  can be approximated as

$$p(\pi_t | \mathbf{Y}^t) \approx \sum_{j=1}^M w_t^{(j)} \delta(\pi_t - \pi_t^{(j)}). \quad (14)$$

Substituting (14) into (8) yields

$$p(\Psi_t | \mathbf{Y}^t) = \sum_{j=1}^M w_t^{(j)} p(\Theta_t | \pi_t^{(j)}, \mathbf{Y}^t) \delta(\pi_t - \pi_t^{(j)}) \quad (15)$$

Now, we have to sample from the space of  $\pi_t$ . The weights can be evaluated recursively:

$$w_t^{(j)} \propto \frac{p(\mathbf{y}_t | \pi_t^{(j)}) p(\pi_t^{(j)} | \pi_{t-1}^{(j)})}{f(\pi_t^{(j)} | \pi_{t-1}^{(j)}, \mathbf{y}_t)} w_{t-1}^{(j)}. \quad (16)$$

For exact marginalization, all proofs of global convergence hold, Doucet *et al*, 2000.

5. IMPLEMENTATION Let the model (1) be a linear  $N$ th-order autoregressive model with Gaussian disturbances

$$p(y_t | \mathbf{Y}^{t-1}, \Theta_t) = p(y_t | \varphi_t, \theta_t, \mathbf{Y}^{t-1}) \sim \mathcal{N}(\varphi_t' \theta_t, \sigma^2), \quad (17)$$

where  $\varphi_t = (y_{t-1}, \dots, y_{t-N}, 1)'$  is a regression vector and  $\theta_t \in \mathbb{R}^{N+1}$  is a vector of regression coefficients, and  $\sigma^2 \in \mathbb{R}^+$  is the noise variance. The Bayesian paradigm exploits the Gauss-inverse-Wishart distribution as a suitable conjugate prior distribution, Kárný (2006).

$$p(\Theta_t | \mathbf{Y}^{t-1}) \sim \mathcal{GiW}(\mathbf{V}_t, \mathbf{v}_t), \quad (18)$$

where  $\mathbf{V}_t \in \mathbb{R}^{(N+1) \times (N+1)}$  denotes an extended information matrix, i.e., a positive-definite symmetric matrix. The term  $\mathbf{v}_t \in \mathbb{R}^+$  stands for the degrees of freedom, Peterka (1981). The data update rule (2) reads

$$\mathbf{V}_t = \mathbf{V}_{t-1} + (y_t, \varphi_t')' (y_t, \varphi_t')$$

$$\mathbf{v}_t = \mathbf{v}_{t-1} + 1.$$

There are various methods accomplishing the time update based on forgetting (5), e.g. Peterka (1981), Jazwinsky (1970), Kulhavý and Kárný (1984), and many others. The approximation of mixture (6) of  $\mathcal{GiW}$  pdfs, in the sense of minimizing the Kullback-Leibler divergence (7), is thoroughly described in Dedecius (2010). The weights  $\pi_t$  are sampled from the Dirichlet distribution  $\mathcal{Dir}(\pi_t)$  by a particle filter. The evolution model  $\pi_t | \pi_{t-1}$  is given by the following transition pdf

$$\pi_t | \pi_{t-1} \sim \mathcal{Dir}(\pi_{t-1} / \Delta + s), \quad (19)$$

where  $\Delta$  is the width of the random walk and  $s$  is the stabilization term.  $\Delta$ ,  $s$  and the proposal distribution are *a priori* given by the user. A popular choice of the proposal distribution  $f\left(\pi_t^{(j)}|\pi_{t-1}^{(j)}, y_t\right) = p\left(\pi_t^{(j)}|\pi_{t-1}^{(j)}\right)$  simplifies (16),

$$w_t^{(j)} \propto p\left(y_t|\pi_t^{(j)}\right) w_{t-1}^{(j)}.$$

There exist other optimal choices of proposal density as well, see, e.g., Doucet *et al.* (2001).

## 6. SIMULATION

In this simulation, we analyze a time series  $y_t = x(t) + e_t$ , where  $x(t)$  is given by the  $x$ -component of the Lorenz system, Lorenz (1963)

$$\dot{x} = \sigma(y - x)$$

$$\dot{y} = \rho x - y - xz$$

$$\dot{z} = -\beta z + xy$$

and  $e_t \sim \mathcal{N}(0, 1)$ . We numerically integrate the system by the fourth-order Runge-Kutta algorithm with time step 0.05 and parameters  $\sigma = 10$ ,  $\rho = 28$ , and  $\beta = 8/3$ . The integration was initialized with  $x_0 = 0, y_0 = 1$  and  $z_0 = 1.05$ . The sampling period coincides with the integration step. The system was modelled using a second-order autoregressive model (17) with  $\phi_t = (y_{t-1}, y_{t-2}, 1)$ . Its parameters were estimated using partial forgetting with hypotheses formulated in Section , where  $\mathfrak{F}$  is the exponential forgetting with factor 0.95. The result of modelling the first 500 samples is depicted in Fig. 1. We can see, that after the learning period (approximately first 100 samples), the estimator stabilizes and the model achieves a good performance.

## CONCLUSION

The autoregressive model with partial forgetting within the Rao-Blackwellized particle filter was discussed. We presented a hybrid filtering method, where a subset of parameters is estimated using a particle filter. The rest of the parameters are estimated conditionally linear. The presented algorithm in its basic form performs well, however, there is a lot of space for further improvements, e.g., use of an adaptive proposal in the particle filter.

## ACKNOWLEDGEMENT

This research was supported by grants MV ĀR VG20102013018 and SGS 10/099/OHK3/1T/16.

## BIBLIOGRAPHY

- Bernardo, J.M. and Smith, A.F.M. (2001). Bayesian theory. *Measurement Science and Technology*, **12**.
- Dedecius, K. (2010). Partial forgetting in Bayesian estimation. PhD Thesis. Czech Technical University in Prague.
- Doucet, A., De Freitas, N. and Gordon, N. (2001). Sequential Monte Carlo methods in practice. Springer.

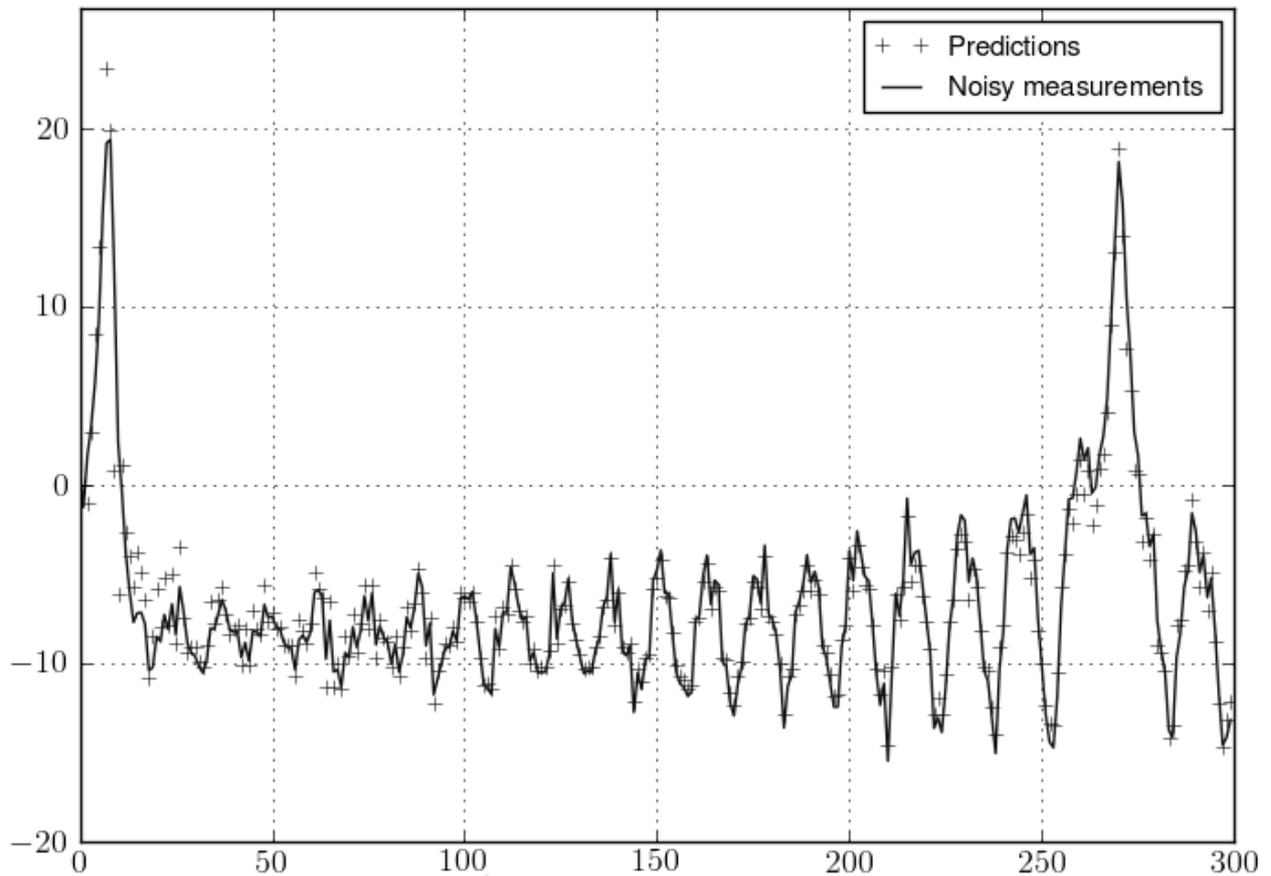


Figure 1: Results of the numerical experiment.

Doucet, A., Godsill, S. and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and computing*, **10(3)**, 197–208

Jazwinski, A.H. (1970). *Stochastic processes and filtering theory*, Academic Press.

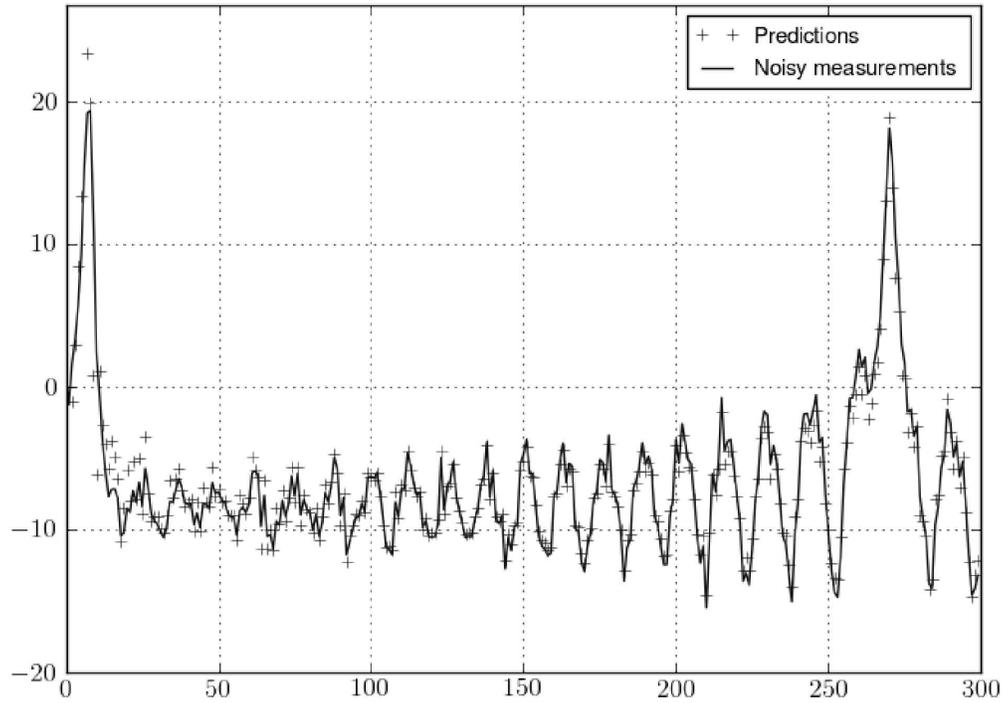
Judd, K. and Mees, A. (1995). On selecting models for nonlinear time series. *Physica D: Nonlinear Phenomena*, **82(4)**, 426–444.

Kalman, R.E. (1957). Physical and mathematical mechanisms of instability in nonlinear automatic control systems. *Trans. ASME*, **79(3)**, 553–566.

Karush, J. (1961). On the Chapman-Kolmogorov equation. *The Annals of Mathematical Statistics*, **32(4)**, 1333–1337.

Kulhavý, R. and Kárný, M. (1984). Tracking of slowly varying parameters by directional forgetting. *9th TFAC World Congress, Budapest, Hungary*, 1984.

- 1 Kullback, S. and Leibler, R.A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 79–86.
- 2
- 3 Kárný, M. (2006). Optimized Bayesian dynamic advising: Theory and algorithms. Springer.
- 4
- 5 Lorenz, N.E. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, **20**, 130–141.
- 6
- 7
- 8 Peterka, V. (1981). Bayesian system identification. *Automatica*, **17**(1), 41–53.
- 9
- 10 Raftery, A.E., M. Kárný, J. Andrýsek, and P. Ettlér. (2010). Online prediction under model uncertainty via dynamic
- 11 model averaging: application to a cold rolling mill. *Technometrics*, **52**(1).
- 12
- 13
- 14 Ragot, J., Mourot, G. and Maquin, D. (2004). Parameter estimation of switching piecewise linear system. *Proceedings*
- 15 *of 42nd IEEE Conference on Decision and Control*, **6**, 5783–5788.
- 16
- 17
- 18 Ristic, B., Arulampalam, S. and Gordon, N. (2004). Beyond the Kalman filter: Particle filters for tracking applications.
- 19 Artech House Publishers.
- 20
- 21
- 22 Rosenqvist, F. and Karlström, A. (2005). Realisation and estimation of piecewise-linear output-error models. *Auto-*
- 23 *matica*, **41**(3), 545–551.
- 24
- 25
- 26 Savona, C. (1988). Approximate nonlinear filtering for piecewise linear systems. *Systems & Control Letters*, **11**(4),
- 27 327–332.
- 28
- 29
- 30 Sontag, E. (2002). Nonlinear regulation: The piecewise linear approach. *Automatic Control, IEEE Transactions on*.
- 31 **26**(2), 346–358.
- 32
- 33
- 34 Tong, H. (1993). Non-linear time series: a dynamical system approach. Oxford University Press.
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60



241x168mm (600 x 600 DPI)

view Only

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60