On particle filters applied to electricity load forecasting

Zakarya ALI, Antoine GRELETY, Samir TANFOUS
19/01/2018

Plan

1. Introduction
2. Model
3. Data
4. Particle Filter
5. PMCMC
6. Results
7. Conclusion

On particle filters applied to electricity load forecasting

Published in 2013
By Tristan Launay, Anne Philippe and Sophie Lamarche

Recent developments regarding:

- Sequential Monte Carlo Methods
- Online prediction
- Wide variety of particle filters applications :
  - Filtering
  - Predicting
- Computationally cheaper approach than classic MCMC algorithms
- As effective as other proven methods

Model

Observed electricity load
At day n
And time i (i = 1...48)

Model state

Noise

$$y_{n,i} = x_{n,i} + v_{n,i}$$

$$v_{n,i} \sim \mathcal{N}(0, \sigma^2)$$

$$x_n = x_n^{season} + x_n^{heat} + x_n^{cool}$$

Not relevant enough
regarding French electricity

$$x_n^{season} = s_n . \kappa_{daytype_n}$$

$$x_n^{heat} = g_n^{heat}(T_n^{heat} - u^{heat})\mathbb{1}_{(u^{heat} > T_n^{heat})}$$

Coefficient depending of daytypes
We chose $K = 1$

$$x_n^{season} = s_n . \boxed{\kappa_{daytype_n}}$$

$$s_n = s_{n-1} + \epsilon_n^s, \text{ with } \epsilon_n^s \sim \mathcal{N}(0, \sigma_{s,n}^2,]-s_{n-1}, +\infty[)$$

$$\sigma_{s,n} = \sigma_{s,n-1} + \eta_n^s, \text{ with } \eta_n^s \sim \mathcal{N}(0, \sigma_s^2,]-\sigma_{s,n-1}, +\infty[)$$

Temperature at time n

Threshold

$$x_n^{heat} = g_n^{heat}(\boxed{T_n^{heat}} - \boxed{u^{heat}})\mathbb{1}_{(u^{heat} > T_n^{heat})}$$

$$g_n^{heat} = g_{n-1}^{heat} + \epsilon_n^g, \text{ with } \epsilon_n^g \sim \mathcal{N}(0, \sigma_{g,n}^2,]-\infty, -g_{n-1}^{heat}[)$$

$$\sigma_{g,n} = \sigma_{g,n-1} + \eta_n^g, \text{ with } \eta_n^g \sim \mathcal{N}(0, \sigma_g^2,]-\sigma_{g,n-1}, +\infty[)$$

Fixed parameters

$$\boxed{\theta = (\sigma, \sigma_s, \sigma_g, u^{heat}, \kappa)}$$

Data

- French electricity load
    - **RTE** : Consumption every 30 min
- Daily average temperatures in French metropolitan territory
    - **Meteo France** : Input every 3 hours
- Day type: weekday, weekend, bank holiday,...
- We chose to work on 2 models : **3AM** and **3PM**

Particle Filter

$$\mathrm{ESS}(n) = \frac{1}{\sum_{k=1}^{M} (w_n^k)^2}.$$

At time $n \geq 1$

1. Sample $\widehat{X}_n^j \sim f_n(x_n | X_{n-1}^j)$.
2. Compute $\widetilde{w}_n^j = w_{n-1}^j g_n(y_n | X_n^j)$ and set $\widehat{w}_n^j \leftarrow \frac{\widetilde{w}_n^j}{\sum_{k=1}^{M} \widetilde{w}_n^k}$.
   - if $\widehat{\mathrm{ESS}}(n) < 0.001M$, set $X_n^j \leftarrow \widehat{X}_n^j$ and $w_n^j \leftarrow w_{n-1}^j$.
   - if $0.001M \leq \widehat{\mathrm{ESS}}(n) < 0.5M$, use residual-multinomial resample (see Algorithm 3.5) and regularisation move (see Algorithm 3.6) steps to set X_n^j and w_n^j.
   - if $0.5M \leq \widehat{\mathrm{ESS}}(n)$, set $X_n^j \leftarrow \widehat{X}_n^j$ and $w_n^j \leftarrow \widehat{w}_n^j$.

1. Resampling at n=0
2. Forecast x_n | x_n-1
3. Generate M new weights
4. Resampling:
   a. If ESS critically low consider observation as missing
   b. Multinomial resampling if ESS is too low
   c. Otherwise keep the new weights
5. Move to n+1

Particle Filter - Initialization

3 methods to initialize the particle filter :

- Hand-picked values
- Use article priors

$$\sigma^2_{s,*}, \sigma^2_{g,*} \sim \mathcal{IG}(10^{-2}, 10^{-2}) \qquad s_0 \sim \mathcal{N}(0, 10^8, \mathbb{R}_+) \qquad g_0^{\text{heat}} \sim \mathcal{N}(0, 10^8, \mathbb{R}_-)$$

- Gibbs sampler
    - Compute the full conditionals of the the smoothed distribution

$$\pi\left(s_{0:n_0-1}, g^{heat}_{0:n_0-1}, \sigma_{s,*}, \sigma_{g,*} \middle| y_{0:n_0-1}\right)$$

Particle Filter - Resampling

- How to fight particles degeneracy ?
- Particles with important weight must remain, while we replace low weight particles

Multinomial resampling :

$$\text{At time } n \geq 0$$
$$\text{Sample } Z_n^j \sim \sum_{n=1}^{M} \omega_n^k \delta(X_n^k, dx)$$
$$\text{Replace } X_n^j \leftarrow Z_n^j \text{ and } \omega_n^k \leftarrow 1/M$$

PMMH

- Estimation of the parameters $\Theta = (\mu^{heat}, \sigma^2_{s,0}, \sigma^2_{g,0}, \sigma^2)$
- Define proposal density and hyperparameters
- Apply Metropolis-Hastings algorithm
  - Generate new proposal parameters and run the particle filter to compute the likelihood
  - Compute joint prior density of parameters
  - Compute independent gaussian proposal h()

GIMH (Beaumont, 2003)

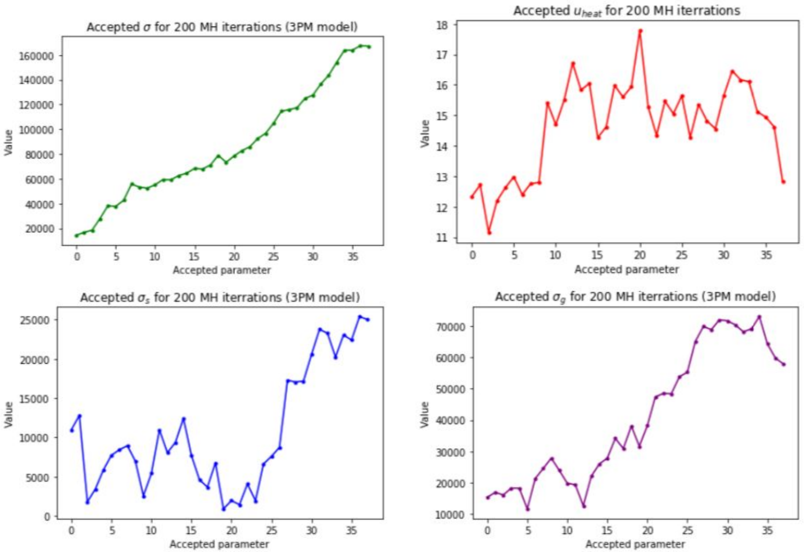
From current point θ_m

1. Sample $\theta_\star \sim H(\theta_m, \mathrm{d}\theta_\star)$
2. With prob. $1 \wedge r$, take $\theta_{m+1} = \theta_\star$, otherwise $\theta_{m+1} = \theta_m$, with

$$r = \frac{p(\theta_\star)\hat{p}(y|\theta_\star)h(\theta_m|\theta_\star)}{p(\theta_m)\hat{p}(y|\theta_m)h(\theta_\star|\theta_m)}$$

Results - PMMH

- **Goal:** Estimate parameters of particle filter for 3am and 3pm
  - 200 iterations
  - Run the particle filter over 715 days
  - 4 hours running time
- Variances are higher than expected
- Problem: likelihood estimated by the particle filter is extremely low



Acceptance rate	μ_{heat}	σ_s	σ_g	σ
19%	12.8	$5.7.10^4$	$2.5.10^4$	$1.6.10^5$

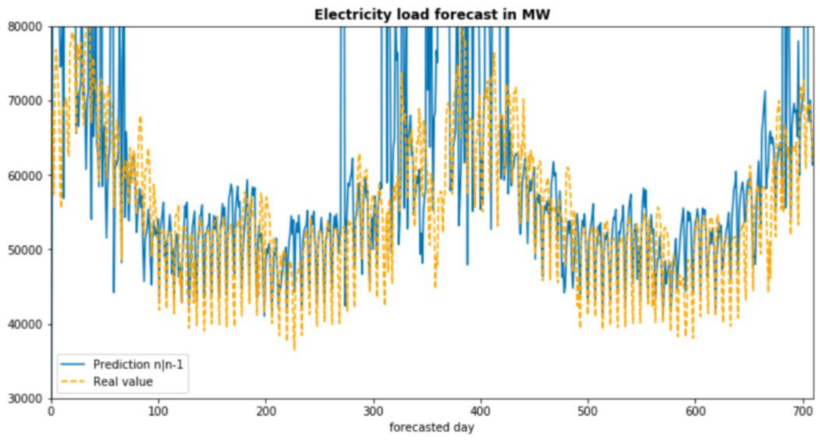
Table 2 : Estimated parameters PMMH 3pm

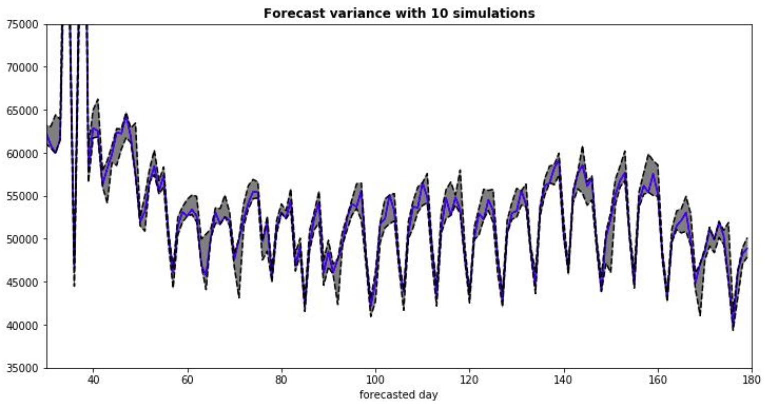
Acceptance rate	μ_{heat}	σ_s	σ_g	σ
29%	8.1	$1.8.10^4$	$8.4.10^4$	$2.9.10^5$

Table 1 : Estimated parameters PMMH 3am

Results - Forecast

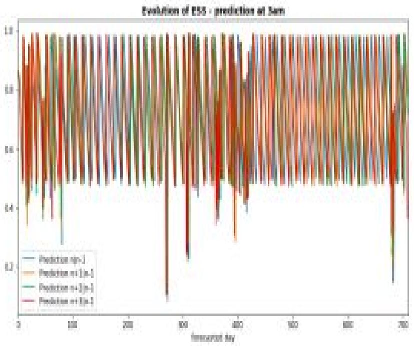
- Online prediction of electricity load 1 to 4 days ahead over 715 days
- Run particle filter initialized with Gibbs and parameters set by hand
- Predicted value +/- 25% distant from the real value
- Outlier values during Christmas period

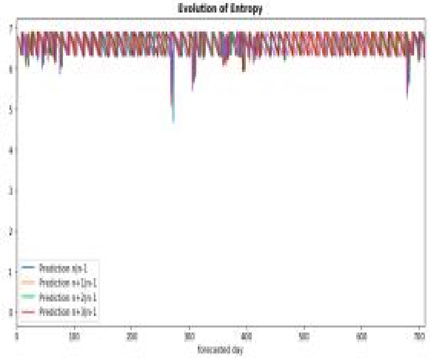


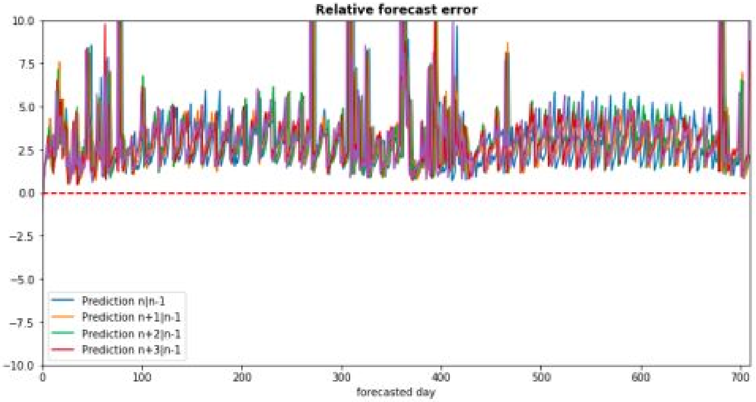


Results - Forecast

- Run particle filter with parameters simulated via PMMH
- Forecasted value about 3 times the real value in average
- Problem: PMMH estimates large variances







Conclusion

- Particle Filter :
  - **Highly sensitive** to initial values
  - Remains computationally intensive
- Our PMMH calibration is **unusable**
- **However** much better forecasting with manual parameters
- Ways to improve :
  - Increase the number of particles for theta estimation / initialization steps / forecasted days...
  - Try theta estimation with Robbins-Monro

Questions?

Robbins-Monro

- Other method to initialize the parameters μ_{heat} , σ , σ_g and σ_s .
- Maximizing method for noisy function

We estimate parameters as follow :

$$\theta_{n+1} = \theta_n - a_n(N(\theta_n) - \alpha) \quad \text{with} \quad \sum_{n=0}^{\infty} a_n = \infty \quad \text{and} \quad \sum_{n=0}^{\infty} a_n^2 < \infty$$

With N the estimated log-likelihood gradient