7th International Building Physics Conference

IBPC2018

Proceedings

SYRACUSE, NY, USA

September 23 - 26, 2018

Healthy, Intelligent and Resilient
Buildings and Urban Environments

ibpc2018.org | #ibpc2018

Sequential Monte Carlo for on-line estimation of the heat loss coefficient

Simon Rouchier[1,*] Maria José Jiménez[2], Sergio Castaño[2]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LOCIE, 73000 Chambéry, France
[2]CIEMAT, Department of Energy, Energy Efficiency in Buildings Unit, Av. Complutense 22, E28040 Madrid, Spain

[*]*Corresponding email: simon.rouchier@univ-smb.fr*

ABSTRACT
The calibration of building energy models based on in-situ sensor information is generally performed after the measurement period, using all data in a single batch. Alternatively, on-line parameter estimation proposes updating a model every time a new data point is available: this allows observing a direct relation between external events and the identifiability of parameters. The present study uses the Sequential Monte Carlo method to train a RC model, and thus estimate a Heat Loss Coefficient, and other parameters, sequentially. Results show the direct impact of solicitations (solar irradiance and indoor heat input) on this estimation, in real time. The method is validated by comparing its results with the Metropolis-Hastings algorithm for off-line estimation.

KEYWORDS
Bayesian inference; on-line; SMC

INTRODUCTION
The calibration of simplified building energy models using in-situ measurements is now a widespread research topic (Rouchier, 2018). It is commonly performed for two general types of applications: the characterisation of intrinsic building performance (Heo et al. 2012; Bauwens and Roels, 2014), and the identification of a model for predictive purposes, for instance in the aim of model predictive control. State-space models, which include the simplified resistor-capacitor (RC) model structures, are a popular choice for both applications. When written as a set of Stochastic Differential Equations, they allow accounting for modelling approximations (Madsen and Holst, 1995) and offer a more reliable parameter estimation than deterministic models (Rouchier et al. 2018).

Parameter estimation is typically performed *off-line*: measurements of indoor and outdoor conditions are first carried in a test building, and data is processed after the experiment in a single batch. An interesting challenge is to carry parameter estimation *on-line*, during the observation period: starting from an initial guess for parameter values, these estimates are updated sequentially, every time a new observation becomes available. The motivation is twofold: first, it would allow using the measurement period for computations, thus reducing the total time of the procedure (Raillon and Ghiaus, 2017). Second and foremost, it would allow directly observing which phenomena “bring information” to the parameters, by correlating the reduction in their estimation uncertainty with observed events.

Bayesian inference offers the possibility of on-line estimation with Sequential Monte-Carlo (SMC) methods (Doucet et al. 2000). Originally developed for the sequential estimation of states (Handschin 1970), SMC was later adapted to state and parameter estimation (Kantas et al. 2015). Building physics applications are scarce and very recent (Raillon and Ghiaus, 2017), but may become more common due to the motivations listed above. The present paper applies SMC for the on-line estimation of the heat loss coefficient (HLC) of a test building. Starting from a highly uncertain prior knowledge of HLC, the target is to dynamically observe what leads its estimation to narrow down to a more precise value. The identifiability of HLC regarding available data is then discussed.

CASE STUDY

Test case

This study uses measurements that were carried in the Round Robin Test Box (RRTB), within the framework of the IEA EBC Annex 58 (Jimenez et al. 2016). This experimental test cell has a cubic form with dimensions of one cubic meter, identical wall components on all sides and one window with dimensions 60x60cm^2. It was installed outdoors, in the LECE laboratory at Plataforma Solar de Almeria, in the South East of Spain. Experiments were carried during a 43 days period in the winter of 2013-2014. More information on the conditions of the test is available in the Annex 58 report (Jimenez et al. 2016).

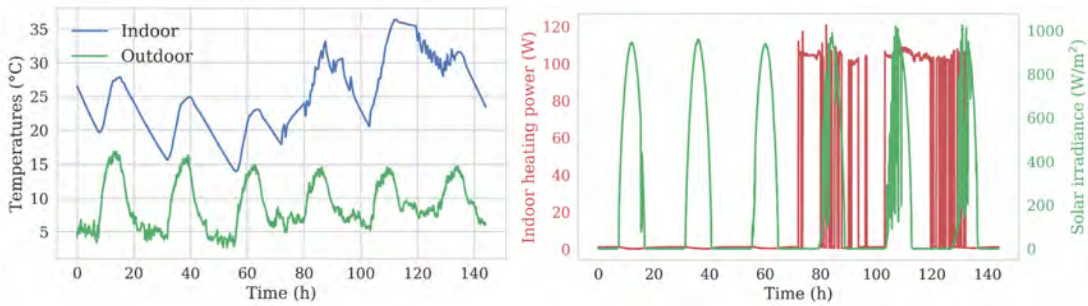


Figure 1. Measured indoor and outdoor temperature, heating power and solar irradiance

A period of six days was chosen for the present investigation, shown on Fig. 1. The indoor temperature is left free floating during the first three days, although it is impacted by daily peaks of solar irradiance. Then, a ROLBS heating signal was imposed inside the test box during the remaining three days. The motivation behind the choice for this sequence of measurements is that these boundary conditions are not very informative at first, then become more informative: we expect to witness their effects on the evolution of the estimation of the heat loss coefficient.

Model

The test box is represented by a 2R2C model described by:

$$\begin{bmatrix} \dot{T}_e(t) \\ \dot{T}_i(t) \end{bmatrix} = \begin{bmatrix} \frac{-1}{R_1 C_1} - \frac{1}{R_2 C_1} & \frac{1}{R_2 C_1} \\ \frac{1}{R_2 C_2} & \frac{-1}{R_2 C_2} \end{bmatrix} \begin{bmatrix} T_e(t) \\ T_i(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{R_1 C_1} & 0 & \frac{k_1}{C_1} \\ 0 & \frac{1}{C_2} & \frac{k_2}{C_2} \end{bmatrix} \begin{bmatrix} T_a(t) \\ q(t) \\ I_{sol}(t) \end{bmatrix} + w(t) \tag{1}$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} T_e(t) \\ T_i(t) \end{bmatrix} + v(t) \tag{2}$$

where T_i, T_a and T_e are the indoor, ambient (outdoor) and envelope temperatures. The model has two states T_e (unobserved) and T_i (observed); q (W) is the indoor heating power; I_{sol} (W/m2) is the solar irradiance on a southern vertical plane; R_1 and R_2 (m^2.K/W) are two thermal resistances, C_1 and C_2 (J/K) are thermal capacities, and k_1 and k_2 (m^2) are two solar aperture coefficients, one for each state of the model. $\mathbf{w}(t)$ denotes a Wiener process that represents modelling errors with an incremental covariance $\mathbf{Q}_c$ (Madsen and Holst, 1995), and $\mathbf{v}(t)$ is the measurement error of the indoor temperature, normally distributed white noise with zero mean and variance $\mathbf{R}_c$. Both are considered unknown and will be estimated along with the other static parameters of the model, which are all denoted by a single vector θ.

The stochastic model described by Eq. 1 and 2 must be discretized to specify its evolution between discrete time coordinates.

$$x_t = F\, x_{t-1} + G u_t + w_t \tag{3}$$
$$y_t = H\, x_t + v_t \tag{4}$$

where $\mathbf{x}_t$ denotes the vector of states at the time coordinate t, and $\mathbf{y}_t$ denotes the observations. The reader is referred to (Madsen and Holst, 1995) and (Rouchier 2018) for more details regarding the discretization steps.

Given a state transition probability $p\left(x_t \vee \theta, x_{t-1}, u_t\right)$ (Eq. 3) and an observation probability $p\left(y_t \vee x_t\right)$ (Eq. 4), filtering produces $p\left(x_t \vee y_{1:T}, \theta\right)$, the probability distribution function of each state given measurements and parameter values, and the marginal likelihood function $L_y(\theta) = p\left(y_{1:T} \vee \theta\right)$. The Kalman filter algorithm is not described here for the sake of concision, but has been described by many authors including (Madsen and Holst, 1995; Rouchier 2018).

BAYESIAN PARAMETER ESTIMATION

The target of the on-line parameter estimation exercise is to assess the value of all static parameters of the model, at each time coordinate of the measurement period: the expected output is a sequence of posterior distributions $\left\{p\left(\theta \vee y_{1:t}\right), t \in 1 \ldots T\right\}$, where T is the number of data points in the measurement period. This sequential estimation is performed by the SMC algorithm. For the sake of validation, the estimation has also been carried in a “traditional” off-line fashion with the Metropolis-Hastings algorithm. Both methods are described below.

Off-line estimation: Marginal Metropolis Hastings

The Marginal Metropolis Hastings (MMH) algorithm is part of the family of Markov Chain Monte Carlo (MCMC) methods. Given one batch of data $y_{1:T}$, a parameter prior $p(\theta)$ and a model specification (Eq. 3 and 4), the MMH algorithm returns a finite sequence of samples $\left\{\theta_n, n \in 1 \ldots N\right\}$ approximating the posterior distribution $p\left(\theta \vee y_{1:T}\right)$. Examples of applications to the calibration of building energy models include (Heo et al. 2012; Rouchier et al. 2018).

The algorithm employs a Kalman filter to compute the states $p\left(x_{1:T} \vee \theta, y_{1:T}\right)$ and likelihood $L_y(\theta)$ associated to each proposal for θ. If the state-space model (Eq. 3) is non-linear, this filter can be replaced by a particle filter: this approach is known as Particle Markov Chain Monte Carlo (PMCMC).

On-line estimation: Sequential Monte Carlo

The SMC algorithm for parameter estimation is an adaptation of particle filtering for state variables. The foundation of this method is the Importance Sampling paradigm as described by (Cappé et al. 2007): simulating samples under an instrumental distribution and then approximating the target distributions by weighting these samples using appropriately defined importance weights. The reader is referred to (Cappé et al. 2007) and (Kantas et al. 2015) for a deeper explanation of SMC and its application to parameter estimation. The method used here is inspired from the Iterated Batch Importance Sampling algorithm (Chopin 2002). It is described on Fig. 2.

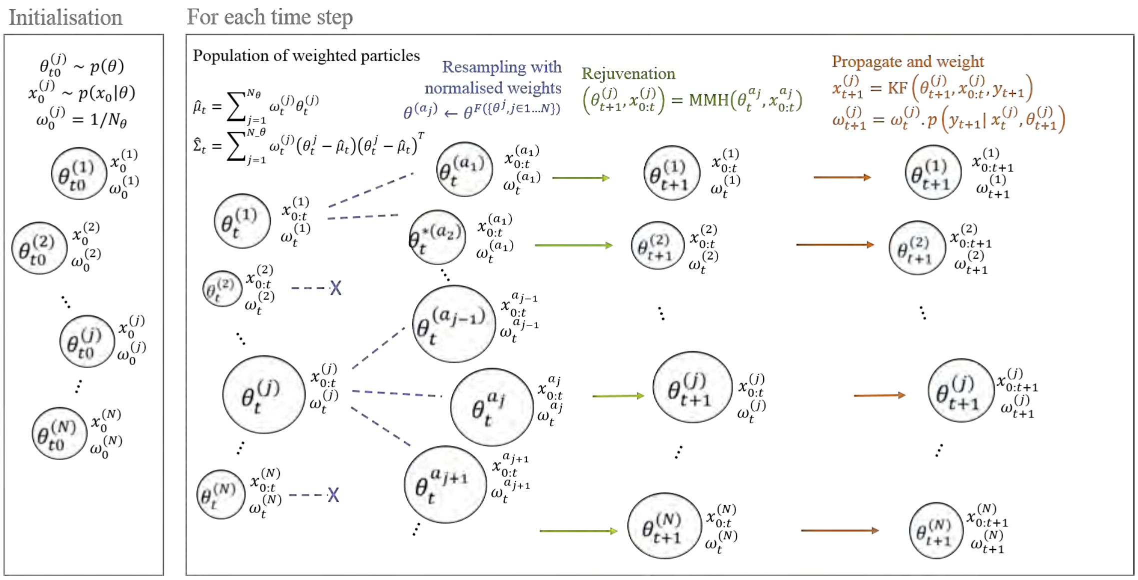


Figure 2. Principle of the Sequential Monte Carlo algorithm

The algorithm starts with the generation of a population of N_θ particles drawn from a prior distribution $p(\theta)$. Each parameter is assigned an initial state $\mathbf{x}_0$ and weight. At each time step t, a Kalman filter computes the states $x_t^{(j)}$ and likelihood $L_t^{(j)}$ associated to each particle $\theta_t^{(j)}$. By this operation, the population of particles is updated so that at each time t they are a properly weighted sample from $p(\theta \vee y_{1:t})$. After several time steps, there is a risk that only a few of the initial particles are significantly more likely than the others and concentrate most of the total weight: a resampling step is then performed to generate a new population of particles from the most influential ones, and a MCMC rejuvenation step then restores the diversity of particles (Murray 2013).

Resampling does not occur every time a new observation becomes available, but only when required: this is measured by the effective number of particles that significantly contribute to the total weight of all particles (Murray 2013). This operation decreases the number of unique particles, hence the subsequent MCMC rejuvenation step that restores diversity. The choice of the proposal distribution for the MCMC rejuvenation step was proposed by (Chopin 2002) and ensures a reasonable acceptance ratio while leaving $p(\theta \vee y_{1:t})$ invariant. The rejuvenation step makes the algorithm quite computationally expensive, since the total likelihood of all particles must be recalculated every time resampling occurs. This problem is mitigated by the fact that particles can be resampled independently, making this effort parallelizable.

RESULTS

The MMH and SMC algorithms were used to estimate the parameter vector θ, either off-line or on-line, using the batch of measurements shown on Fig. 1 from the RRTB test cell represented by a 2R2C model (Eq. 1). First, the sufficiency of the 2R2C model to recreate the indoor temperature time series in checked. This is ensured by the autocorrelation function of the residuals after parameter estimation shown on Fig. 3.

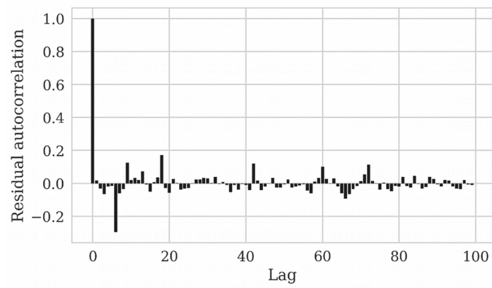
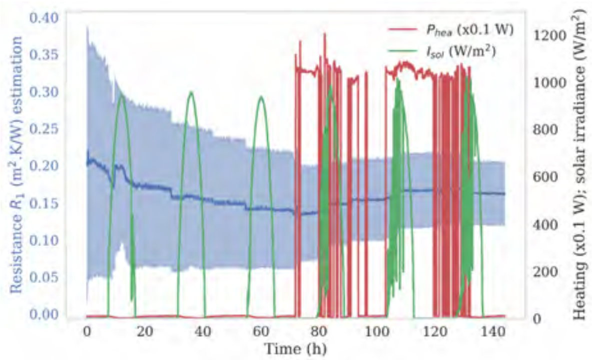
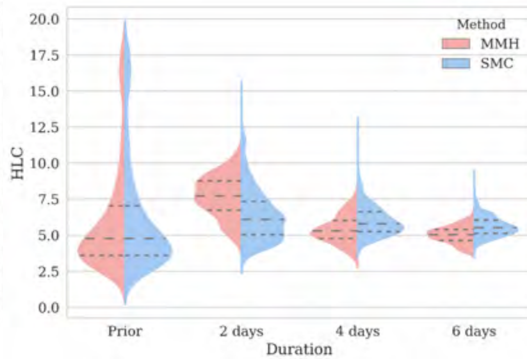


Figure 3. Autocorrelation of one-step prediction residuals

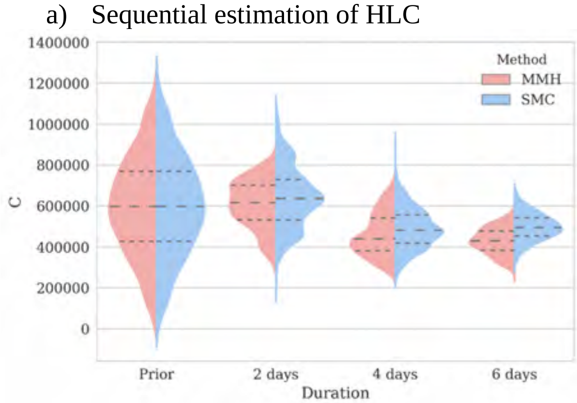
As pointed out by (Madsen and Holst, 1995), the low value of this function ensures that residuals are uncorrelated and thus are close to white noise, in accordance with the hypothesis is the model. This observation allows us to analyze the parameter estimation results. All estimation results from both MMH and SMC are assembled into Fig. 4.



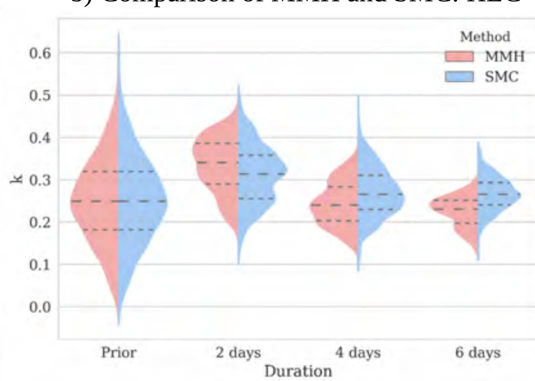
a) Sequential estimation of HLC



b) Comparison of MMH and SMC: HLC



c) Comparison of MMH and SMC: *C*



d) Comparison of MMH and SMC: *k*

Figure 4. Parameter identification results

First, Fig. 4(a) shows the progression of the estimation of the parameter R_1 by SMC during the six days of measurements. The blue area denotes the mean and the 95% confidence region of R_1 at each time coordinate. The indoor heating power and solar irradiance are superimposed over it and allows the interpretation of the estimation. During the first three days of measurements, the value of R_1 can be seen to narrow down to a more confident estimation during the day, when the solar irradiance is positive. An important uncertainty however remains on this parameter. The heating is then turned on inside the test box, which causes the estimation to narrow down more abruptly. These results are consistent with the general knowledge that some indoor solicitations are necessary to make a building energy model identifiable. The originality of the present study is that we can observe the information brought to the parameter estimates in real time.

The performance of SMC is then compared to the MMH algorithm, which estimates parameters from a single batch of data at once. Fig. 4(a), 4(b) and 4(c) respectively show estimates of the heat loss coefficient $HLC=1/(R_1+R_2)$, the total heat capacity $C=C_1+C_2$ and the total solar aperture $k=k_1+k_2$. Each plot shows the prior distribution of these properties, and their estimation after 2 days, 4 days and 6 days of measurements. The MMH algorithm had to be run separately with each measurement length, while the SMC algorithm only had to run once to return all results. Starting from vague Gaussian prior distributions for each parameter, both methods resulted in mostly matching results.

CONCLUSION

The present study uses the Sequential Monte Carlo method to train a RC model, and thus estimate a Heat Loss Coefficient, and other parameters, sequentially. Results show the direct impact of solicitations (solar irradiance and indoor heat input) on this estimation, in real time. The method is validated by comparing its results with the Metropolis-Hastings algorithm for off-line estimation.

ACKNOWLEDGEMENTS

The author would like to thank the French National Research Agency (ANR) for funding this work through the BAYREB research project (ANR-15-CE22-0003).

REFERENCES

Bauwens G., Roels S. 2014. Co-heating test: a state-of-the-art. *Energy and Buildings*, 82, 163-172

Cappé O., Godsill S. and Moulines E. 2007. An overview of existing methods and recent advances in Sequential Monte Carlo. In: *Proceedings of the IEEE*, 95 (5), 899-924

Chopin N. 2002. A sequential particle filter method for static models. *Biometrika*, 89, 539-552

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

Handschin J.E. 1970. Monte Carlo techniques for prediction and filtering of non-linear stochastic processes. *Automatica*, 6(4), 555-563

Heo Y., Choudhary R. and Augenbroe G.A. 2012. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings*, 47, 550-560

Jimenez M.J. et al. 2016. Report of subtask 3a: Thermal performance characterization based on full scale testing - description of the common exercises and physical guidelines. In: *IEA EBC Annex 58*

Kantas N., Doucet A., Singh S.S., Maciejowski J. and Chopin N. 2015. On particle methods for parameter estimation in state-space models. *Statistical Science*, 30(3), 318-351

Madsen H. and Holst J. 1995. Estimation of continuous-time models for the heat dynamics of a building. *Energy and Buildings*, 22, 67-79

Murray, L. M. 2013. Bayesian state-space modelling on high-performance hardware using LibBi. *arXiv preprint arXiv:1306.3277*.

Raillon L. and Ghiaus C. 2017 Sequential Monte-Carlo for states and parameter estimation in dynamic thermal models. In: *Proceedings of the Building Simulation 2017 conference*, San Francisco, USA

Rouchier S. 2018. Solving inverse problems in building physics: an overview of guidelines for a careful and optimal use of data. *Energy and Buildings*, 166, 178-195

Rouchier S., Rabouille M., Oberlé P. 2018. Calibration of simplified building energy models for parameter estimation and forecasting: stochastic versus deterministic modelling. *Building and Environment*, 134, 181-190