Kuzmanovski, Vladimir; Hollmen, Jaakko

Composite Surrogate for Likelihood-Free Bayesian Optimisation in High-Dimensional Settings of Activity-Based Transportation Models

Published in:
Advances in Intelligent Data Analysis XIX - 19th International Symposium on Intelligent Data Analysis, IDA 2021, Proceedings

DOI:
10.1007/978-3-030-74251-5_14

Published: 01/01/2021

Document Version
Peer reviewed version

Composite surrogate for likelihood-free Bayesian optimisation in high-dimensional settings of activity-based transportation models*

Vladimir Kuzmanovski[1,2,4] [0000−0001−7355−1581] and Jaakko Hollmén[1,3]

[1] Department of Computer Science, Aalto University, Finland
{vladimir.kuzmanovski,jaakko.hollmen}@aalto.fi
[2] Department of Knowledge Technologies, Jožef Stefan Institute, Slovenia
vladimir.kuzmanovski@ijs.si
[3] Department of Computer and Systems Sciences, Stockholm University, Sweden
jaakko.hollmen@dsv.su.se
[4] Smart City Center of Excellence, TalTech, Estonia

Abstract. Activity-based transportation models simulate demand and supply as a complex system and therefore large set of parameters need to be adjusted. One such model is Preday activity-based model that requires adjusting a large set of parameters for its calibration on new urban environments. Hence, the calibration process is time demanding, and due to costly simulations, various optimisation methods with dimensionality reduction and stochastic approximation are adopted. This study adopts Bayesian Optimisation for Likelihood-free Inference (BOLFI) method for calibrating the Preday activity-based model to a new urban area. Unlike the traditional variant of the method that uses Gaussian Process as a surrogate model for approximating the likelihood function through modelling discrepancy, we apply a composite surrogate model that encompasses Random Forest surrogate model for modelling the discrepancy and Gaussian Mixture Model for estimating the its density. The results show that the proposed method benefits the extension and improves the general applicability to high-dimensional settings without losing the efficiency of the Bayesian Optimisation in sampling new samples towards the global optima.

Keywords: transportation model · high-dimensional data · Bayesian optimisation · likelihood-free inference · random forest

1 Introduction

Activity-based transportation models (ABM) are designed to simulate the transportation demand and supply as a self-organising agent-based complex system [11]. Such models simulate activities per agent or individual, resulting in a costly execution in terms of computational time. In addition, the transportation-related

*This work has been supported by the European Commission through the H2020 project Finest Twins (grant No. 856602)

activities of individuals rely on complex decisions, modelling of which require a large set of parameter adjustments. One such model is Preday ABM, which motivates this study, and plays an important role in a simulation toolset on transportation - SimMobility [1].

Application of the Preday ABM in simulating various environments requires systematic adjustments or calibration of a large set of parameters (further referred to as ABM parameters), in order to align the associated outputs more closely to the observed values or true output statistics. For that purpose, various optimisation methods are adopted, including primarily gradient-free metaheuristics [28, 29, 31]. However, Bayesian inference with the recent developments provides a valuable analytical approach for the calibration process [35, 36], a great advantage of which is the elimination of necessity to simulate a large sample set in finding the global optima [14, 18, 20, 36].

The Bayesian Optimisation for Likelihood-free Inference (BOLFI) [14] is a method for inferring parameters of simulation-based models by modelling the discrepancy between observed and simulated output statistics. Its state-of-the-art performance are achieved with nonparametric approximation of the likelihood function with regression by Gaussian Processes (GPs) [3], applied in various domains, such as population genetics [18], spreading of pathogens [22], atomic structure of materials [40], as well as cosmology [20].

The BOLFI facilitates likelihood-free inference of the response function that maps the parameters' values with the discrepancy in the output statistics, by combining probabilistic (Bayesian) inference and iterative search (optimisation). The former inherits the theory of approximate Bayesian computation (ABC) [24, 22] to support the likelihood-free inference. The iterative search is used for acquiring new samples (parameters' values) that have great potential to direct the search towards global optima (minimum discrepancy), by utilising identified and evaluated optimal points from previous iterations.

However, the ABC methods have limited applicability in settings of high-dimensional data and costly simulations [27, 16, 32], constituting a bottleneck for their broader adoption in settings of complex simulation models. Therefore, applicability of the BOLFI method for calibrating the Preday ABM to a new urban environment is limited, and neither of the proposed solutions, such as dimensionality reduction [7, 35], or introduction of synthetic parametric/nonparametric likelihoods [42, 37, 30, 32, 2], circumvent the obstacle. Namely, the former requires increased number of simulations, while the latter are applicable to problems with low number of parameters.

The aim of this study is to overcome the aforementioned bottleneck, for which we propose an improvement of the BOLFI method with so called Composite Surrogate Model (BOLFIw/CSM) as a robust surrogate model that handles the high-dimensional data and takes advantage of the BOLFI method to limit the number of costly simulations. Namely, the BOLFI method, as specified in [9, 14], uses a Gaussian Process (GP) as a surrogate model for modelling the response, i.e., discrepancy, which limits the applicability on high-dimensional data. On the other hand, if the GP is replaced with more robust surrogate regression model

(e.g., Random Forest as in [15]), then the non-probabilistic point estimates of the posterior affect the acquisition of new samples (through an acquisition function) to efficiently identify the regions of interest in the parameter space, i.e. the exploitation-exploration trade-off [17]. Therefore, the proposed BOLFIw/CSM method adopt the Random Forest [8] as a surrogate model and combines it with a Gaussian Mixture Model [33] as conditional density estimator for semi-parametric estimation of the posterior distribution.

The rest of the manuscript is organised as follows. In Section 2 we introduce the Preday ABM and the BOLFI method, followed by a formalisation of the proposed extension to the BOLFI method, in Section 3. Section 4 introduces the experimental design and discusses the achieved results. Finally, the work and conclusions are summarized in Section 5.

2 Materials and Methods

2.1 Preday ABM

The Preday ABM model is a fundamental part of a comprehensive simulation tool SimMobility and it is used to simulate the mid-term demand of a transportation network in a given urban area [1]. The demand is formulated as daily travel activities of households and individuals, and the simulation is based on population characteristics of the simulated urban area.

The model consists of 22 discrete choice sub-models, with total of 817 parameters. The daily activity schedule of agents are modelled with application of hierarchical discrete choice models using a Monte-Carlo simulation, over predefined set of activities per agent by using the random utility theory [23, 5].

Overall, the daily simulated activities are categorized in accordance to the activity type (*work*, *education*, *shopping*, and *others*) and transportation modes (*MRT*, *Bus*,*Private Bus*, *Car-drive alone*, *Shared car-drive with 2 passengers*, *Shared car-drive with 3 passengers*, *Motorcycle*, *Taxi*, and *Walk*). These categorizations are used for statistical comparison of the simulated data with the observed data, where the output statistics is expressed as a number of activities per different combination of activity type and transportation mode (e.g., number of bus rides for work-based commuting).

Calibration of Preday ABM has been performed in several studies, whereby parameters are estimated by Simultaneous perturbation stochastic approximation (SPSA) method with its variants [28, 29, 31]. However, all studies considered reduction of the dimensionality of the parameter space either by sensitivity analysis (SA) [34] or principal component analysis (PCA) [19].

2.2 Bayesian optimisation for likelihood-free inference

Likelihood-free inference approach is a method for inferring a posterior distribution of parameters of a simulation-based model [14]. The simulation-based model is defined as a generative process that under certain parameter values

generates data similar to observations of an underlying phenomenon. Often the simulation-based models are of black-box nature or unknown analytical form, hence their likelihood function is intractable.

From the Bayesian perspective, the inference task corresponds to a statistical inference of a finite number of parameters $\theta \in \mathbb{R}^d$ of the simulation-based model from a set of observations Y_o:

$$p(\theta|Y_o) = \frac{p(Y_o|\theta) \cdot p(\theta)}{p(Y_o)}, \tag{1}$$

where $p(\theta)$ encodes our prior beliefs on the distribution of parameter values and $p(Y_o|\theta)$ represents the likelihood of the observations, given the parameters, derived from the known function $\mathcal{L}(\theta)$. Since the analytical form of $\mathcal{L}(\theta)$ is unknown in the underlying challenge, we use the notation $L(\theta)$ that need to be approximated over a set of N samples - $\tilde{L}^N(\theta)$. The notation is simplified if the marginal distribution $p(Y_o)$ is omitted because it does not depend on θ:

$$p(\theta|Y_o) \propto L(\theta) \cdot p(\theta), \tag{2}$$

where the $L(\theta)$ is approximated over finite sample set ($\tilde{L}^N(\theta)$) and it is reconstructed as the number of samples increases:

$$\lim_{N\to\infty} \tilde{L}^N(\theta) = L(\theta) \tag{3}$$

The approximation ($\tilde{L}^N(\theta)$) of the likelihood function ($L(\theta)$) can be performed in parametric or non-parametric manner. The former assumes that the likelihood belongs to a certain parametric family, and hence it is called synthetic likelihood [42]. The latter, alternatively, approximate the likelihood function by a kernel density estimation [12] or surrogate regression [14].

The Bayesian Optimisation for Likelihood-free Inference (BOLFI), as an iterative process, approximates the likelihood function from the posterior distribution of the response function, i.e., modelled discrepancy with surrogate regression model, which is updated at each iteration following the Bayes's theorem [14]. The iterative update of the posterior distribution triggers acquiring new evidence from the parameter space with a highest potential to progress towards global optima. So, an *acquisition function* $A(\theta)$ is introduced, whereby $s \in \mathbb{R}$ generated samples are credited with an utility. The iterative process continues by enriching the evidence with simulated $k \leq s$ samples with the highest utility.

In [9, 17] several acquisition functions are defined, but for the purpose of this study, we adopt the *expected improvement* (EI) [26], defined as follows:

$$EI(\theta) = \sigma(\theta)[z\Phi(z) + \phi(z)], \tag{4}$$

$$z = \frac{f^* - \mu(\theta)}{\sigma(\theta)}, \tag{5}$$

where $\sigma(\theta)$ and $\mu(\theta)$ are statistics of the inferred posterior distribution, f^* is the most optimal output, i.e., minimal discrepancy discovered, and Φ and ϕ

are probability density and cumulative distribution function in terms of the standard normal distribution, respectively. The expected improvement $EI(\theta) = 0$ if $\sigma(\theta) = 0$. The analogy behind (4) reveals the exploration-exploitation trade-off that favours larger uncertainty proximal to the known optimal region(s).

2.3 Limitations of BOLFI for calibrating Preday ABM

The Preday ABM can be seen as a black-box model with costly executions, parameters of which need to be adjusted so that the simulated output corresponds to the observed quantities. Therefore, for minimizing the discrepancy in the output with limited number of simulations, the BOLFI method fits naturally. However, its applicability is limited to settings with up to 10-dimensional parametric space [38, 17], which clearly does not suffice the requirement of Preday ABM model with total of 817 ABM parameters.

To overcome the limitation of likelihood-free inference over high-dimensional spaces when applied with Bayesian optimisation, earlier attempts propose sequential investigation of effective sub-spaces [10, 41] or discovery of active subspaces [35]. Both require larger set of simulations and yet have been proven to work for up to 400 features. Recent attempts opt for employment of Deep Gaussian Processes (DGPs) [3] and combination of GPs [39] for splitting the discrepancies in accordance to a latent variable, and dimensions of the parameter space, respectively. Both reported successful application over couple of domains, albeit with much lower order of magnitude in terms of dimensionality, than the underlying problem of calibrating the Preday ABM. Other approaches consider replacement of the regression surrogate model, in particular with Random Forest [15, 32], but on use-cases defined over 76- and 2-dimensional parameter space.

3 BOLFI with Composite Surrogate Model

The proposed extension of the BOLFI method adds on the previous works that consider the Random Forests (RF) [8] as a surrogate model for likelihood approximation through modelling the discrepancy, extended with a Gaussian Mixture model [13, 33] as a density estimator conditioned on the approximated likelihoods. The novelty of the proposed method lies in the density estimation of predictions from a robust surrogate model, which can be utilised by existing acquisition functions that depend on probabilistic inputs. The combination of both components is referred to as a composite surrogate model and the overall proposed method abbreviated as BOLFIw/CSM.

Random Forest [8] is an ensemble method composed of C regression trees. Regression trees follows the concept of a decision tree, with structure made of decision binary nodes, built iteratively in top-down fashion. Each regression tree is built over a sub-space of the parameter space, which is designed by random subsets of both the features (dimensions) and bootstrap samples. Therefore, each regression tree predicts the target, given a dataset, for a specific region in

the defined space. Ensemble prediction, on the other hand, is an aggregation (average) of the outcomes of all C tree components:

$$\mathcal{RF}(\theta|\Theta_o, Y_o) = \frac{1}{C}\sum_{i=1}^{C}\tau_i(\theta|\Theta_{s_i}, Y_{s_i}), \tag{6}$$

where C is the number of tree components, Θ_{s_i} and Y_{s_i} are training dataset of i-th regression tree τ_i, while Θ_o and Y_o global training dataset.

The RF regression method has relatively small number of hyper-parameters that can significantly influence the outcome. Commonly adjusted are: the number of tree components C; the minimum number of samples in a terminating node that controls the structure growth and overfitting settings of the individual tree components; and number of features to design a sub-space or partition.

Gaussian Mixture Model (GMM) [13, 33] is a semiparametric density function composed of weighted sum of M components, where each component is a Gaussian distribution (function). In this study we are considering one-dimensional model over a vector of values $x \in \mathbb{R}^C$, defined as follows:

$$p(x) = \sum_{i=1}^{M} w_i \, \mathcal{N}(x|\mu_i, \sigma_i); \;\; \sum_{i=1}^{M} w_i = 1, \;\; w_i > 0. \tag{7}$$

$$\mathcal{N}(x|\mu_i, \sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} exp(-\frac{1}{2\sigma_i^2}(x-\mu_i)^2), \tag{8}$$

Composite surrogate model (CSM) extends the RF surrogate model with a GMM density estimator. The proposed method takes advantage of the RF as a robust regression method in terms of the high-dimensional data, and compensates its limitations regarding the expected parametric posterior estimates for likelihood approximation. Namely, the RF model is: (i) unable to predict a value outside of the observed range; and (ii) is characterised with the non-probabilistic nature of the predicted outcome. Thus, as a standalone surrogate the RF greatly affects the efficiency of the selected acquisition function $A(\theta)$, i.e., EI, in acquiring new promising samples. Consequently, the GMM extension to the base surrogate model compensate the limitations and estimates the posterior in a semiparametric form from the predicted outcomes. As such, the overall composite nature of the discrepancy surrogate model is considered being semiparametric, as well.

In previous studies where RF surrogate model is used, [15] consider empirical or simple average as posterior point estimates, which retain the limitations, while the study of [32] uses quantile-based estimation of the posterior cumulative distribution by Quantile RF [25]. The latter introduces a probabilistic outcome, but limited to quantile-based representation only, whereby a potential multi-modality of the probability function remains a challenge.

The CSM approximates the likelihood of a given set of parameter values θ as linear combination of component-wise approximated likelihood functions through predicted discrepancies (T) from individual regression trees in a RF:

$$L^N(\theta) = \sum_{i=1}^{M} w_i \; A(\mathcal{N}(T|\mu_i, \sigma_i)) \tag{9}$$

$$T = \{ \; \tau_i \mid \tau_i = \tau_i(\theta|\Theta_{s_i}, Y_{s_i}), \; i = 1, ..., C \; \}. \tag{10}$$

The abstraction behind the proposed method compartmentalises the high-dimensional parameter space into regions examined individually in terms of their exploration-exploitation trade-off. The regions that over the aggregated exploitation points show possibility to improve the exploration will further be favoured.

4 Results

Performances of the proposed method are empirically evaluated through an experimental setup, as described hereafter.

The algorithm of BOLFIw/CSM is described in Algorithm 1. It is initialised with four input parameters: initial and iteration sample size, maximum number of iterations, and random variation rate of new samples. While the rest of the algorithm is covered with previously formalised elements or they are self-sufficiently named, the last parameter and sampling function are to be clarified.

The sampling function is used at the initial phase, when n_i samples are generated and in the iterative process, for generating new n_t samples for evaluation. In case of the former, the function generates random samples (ABM parameter values) from a uniform distribution across all dimensions. For the latter, the sampling function follows the currently optimal set of ABM parameter values and assign new values to a random portion ρ (variation rate) of the dimensions.

Algorithm 1: BOLFIw/CSM

Input:
$n_i/n_t \leftarrow$ initial/iteration sample size;
$max_t \leftarrow$ maximum number of iterations;
$\rho \leftarrow$ variation rate for random variation of new samples;
Result: Θ, S_Θ
$\Theta \leftarrow$ generateInitialSample(n_i);
$S_\Theta \leftarrow$ simulate(Θ);
$\Theta_{best} \leftarrow$ pullOptimalTheta(Θ,S_Θ);
while *! terminate(*max_t*)* **do**
  $\Theta_t \leftarrow$ generateSample(n_t,ρ) ;
  $T \leftarrow$ fitRandomForest(Θ,S_Θ) ;
  $\Lambda_t \leftarrow$ estimateDensity(T,Θ_t) ;
  $EI_t^* \leftarrow$ acquisition(Λ_t) ;
  $\Theta_t^* \leftarrow$ pullOptimalEstimatedTheta(EI_t^*);
  $S_{\Theta_t^*} \leftarrow$ simulate(Θ_t^*);
  $\Theta, S_\Theta \leftarrow$ updateParameterSet(Θ_t^*,$S_{\Theta_t^*}$) ;
end

Table 1. Summary of BOLFIw/CSM and BOLFIw/RF hyper-parameters.

Parameter	BOLFIw/CSM	BOLFIw/RF
Density estimator	GMM [6]	Empirical mean/variance
Variation rate (ρ)	0.3; 0.5; 0.8; 1.0	0.3; 0.5; 0.8; 1.0
Iteration sample size (n_t)	100	100

The newly assigned values are randomly sampled from a Gaussian distribution with unit variance and mean - the value observed in the currently optimal sample, for the corresponding dimensions. Analogously, newly sampled sets of ABM parameter values are generated from the neighbourhood of the optimal ABM parameters. The underlying implementation encompasses RF and GMM algorithms implemented in R and described in [21] and [6], respectively [5].

The data used for empirical evaluation of the proposed method include statistical, economic and demographic description of population of a virtual city, provided along the SimMobility simulation toolbox. The virtual city is designed so that reflects characteristics of the urban area in Singapore, encapsulating population data, built environment, and transportation network description [4]. It is designed with population of 351 518 individuals, 100 000 households, and six modes of transportation (MRT, bus, car, motorcycle, taxi, and walk). In addition, the provided virtual city is specified with calibrated ABM parameter values, which are considered as a ground truth in the discussion of the achieved results within this study.

Due to the time constraints and computationally-demanding simulations, we sampled 10% of the population given in the database. The sampling is performed in stratified manner, according to the features registered in the database.

For comparing purposes, the simulated daily activities are summarised to a set of summary statistics, including: number of tours, number of stops, and number of trips. Furthermore, the summary statistics are calculated in regards to three different categories: (i) per person, (ii) per person and tour type (purpose of travel), and (iii) per person and transportation mode. Therefore, the final simulated results are summarised through total of 9 summary statistics. Finally, the discrepancy between simulations are defined as a Euclidean distance between corresponding vectors of the summary statistics.

The proposed BOLFIw/SCM method is applied for calibrating the Preday ABM to fit the observed daily activities in the virtual city - provided as a supplementary material on the SimMobility transportation simulation toolset. The application is based on predefined set of input parameters of the algorithm, with the variation rate being changed and its effects examined. In addition, for performance evaluating purposes, the experimental setup includes application of the BOLFI method with Random Forest (BOLFIw/RF) as a surrogate model - a method proposed in [15]. Complete set of input parameters concerning the application of both methods are summarised in Table 1. Each run of the both methods with varying variation rate (random portion of the parameters that

[5] Code available at: https://version.aalto.fi/gitlab/kuzmanv1/bolfiwcsm-preday

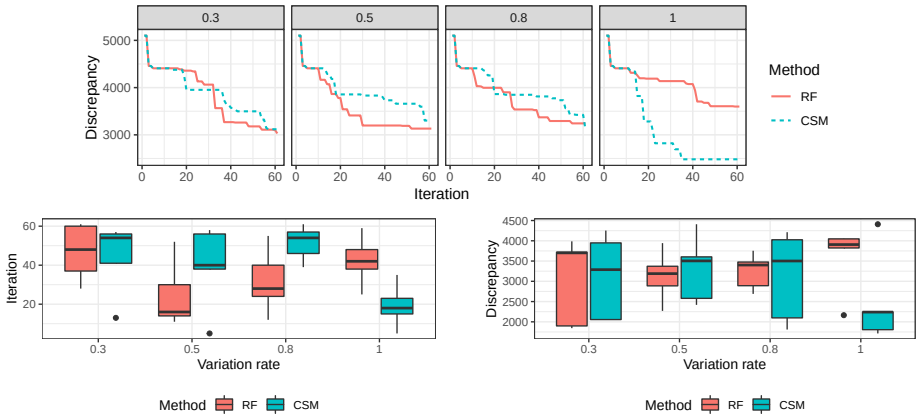

Fig. 1. Average discrepancy reached across all iterations (upper). Variation rates in sampling new parameters compared to the iteration at which the optimal result were first achieved (lower-left) and the discrepancy for both methods (lower-right).

are changed at each sampling) starts with the same initial set of 10 simulated samples and terminated after the 50-th iteration. The Random forest method is applied with default hyper-parameter values (500 regression trees, and minimum number of instances in a leaf is 5. The results are summarised over 5 repeated runs with different random seed (Figure 1).

The results, as shown in Figure 1 (upper figure), distinguish the dominance of the BOLFIw/CSM method, when applied with variation rate of 1.0, which means that sampled parameter values differ completely from the most optimal set at the given iteration. The dominance is observed in terms of pace, at which the final or best score is achieved. Namely, the BOLFIw/CSM method with $\rho = 1.0$ needs approximately 20 iterations on average, for achieving the final results of the other methods (after 50 iterations).

The Figure 1 (lower figures) shows the performance in terms of the iterations when the final optimal result was first achieved, as well as the achieved discrepancy at the end of the iterative processes. The former address the issue of early achievement of the minimum discrepancy, while the latter, the overall optima attained in terms of the discrepancy. In both analyses, when $\rho = 1.0$, the BOLFIw/CSM shows significantly better results with as twice as better performance than the its counterpart. However, aforementioned performances are not attained with lower variation rates, where the progress towards the global minimum is on average similar for both methods, with exception when $\rho = 0.5$.

The analogy behind the observed outcome suggests that the methods with the lower variation rates ($\rho < 1.0$) tend to generate less spread samples in the high-dimensional space, and hence, the density function of the predicted values by the RF surrogate model tends to be more uniformly shaped, without distinct Gaussian components. Therefore, the convergence of the GMM density estimator is affected, resulting in non-reliable density estimates. In such cases, it is apparent

that the empirical average over the predicted values is more informative. On the other hand, the benefit of the GMM density estimator is significant when the sampling function is able to produce scattered new samples, triggering significant progress towards the global optimum.

The calibration of the Preday ABM with the overall lowest discrepancies is achieved by BOLFIw/CSM with variation rate of 1.0, where the distance between the summary statistics derived from simulated and observed data is 1713.26. The best record of its counterpart is 2163.98. Both, however, lags behind the ground truth, whereby the sampled 10% of the population from the virtual city simulates daily activity schedules with discrepancy of 193.89.

5 Summary and Conclusions

This study addresses the issue of calibrating Preday activity-based simulation model from the SimMobility toolset with a limited number of simulations. It corresponds to the task of parameter inference of a simulation-based model with intractable likelihood function over high-dimensional parameter space. For such class of problems, Bayesian Optimisation for Likelihood-free Inference (BOLFI) [14] is of great importance. However, the number of ABM parameters in the Preday ABM is greater than models of previous adoption of the BOLFI method.

Therefore, we aimed at improving the surrogate model to approximate the likelihood through modelling discrepancy. The improved model would be able to encapsulate the knowledge from high-dimensional data with limited sample size, and to tailor the acquisition function to efficiently identify the regions of interest. Consequently, we propose the BOLFI approach with Composite Surrogate Model (BOLFIw/CSM), whereby the posterior distribution for approximating the intractable likelihood function is composed by a density estimates over the regression model. The surrogate model is set to be Random Forest (RF), non-aggregated output of which is fed into a Gaussian Mixture model (GMM) density estimator. The mixture of Gaussians is then used for acquiring new evidence that guides efficiently the search towards the global optima. The proposed method inherits the robust characteristics of the RF models in terms of limited overfit to the small sample size in high-dimensional settings. Moreover, the GMM density estimator adapts the output so that the existing acquisition function can be used without loss of its efficiency in the context of the Bayesian Optimisation.

The BOLFIw/CSM method shows promising results for calibrating the Preday ABM on a new city environment, with data provided as demo data along the SimMobility toolset. The method is compared to the BOLFI approach applied with the RF for approximation of the likelihood, following empirical mean and variance. The BOLFIw/CSM shows great performance with sampling function that generates scattered new samples. Otherwise, as new samples get proximal to the latest most optimal point, the surrogate model tends to predict more uniformly distributed values, which limits the convergence of the density estimator.

In further work, the BOLFIw/CSM method is to be examined in more extensive experimental setup regarding the applicable domains, method's parameters,

and confronting with a larger set of comparative methods. Finally, the calibration of the Preday ABM is to be extended to a real-world urban environment and compared with the non-Bayesian approaches.

References

1. Adnan, M., Pereira, F., Azevedo, C., Basak, K., Lovric, M., et al.: Simmobility: A multi-scale integrated agent-based simulation platform. In: 95th Annual Meeting of the Transportation Research Record (2016)
2. An, Z., Nott, D.J., Drovandi, C.: Robust bayesian synthetic likelihood via a semi-parametric approach. Statistics and Computing **30**(3), 543–557 (2020)
3. Aushev, A., Pesonen, H., Heinonen, M., Corander, J., Kaski, S.: Likelihood-free inference with deep gaussian processes. arXiv preprint arXiv:2006.10571 (2020)
4. Basak, K.: SimMobility Demo Data (2019 (accessed August 1, 2020)), https://github.com/smart-fm/simmobility-prod/wiki/Demo-Data
5. Ben-Akiva, M., Lerman, S.R.: Discrete choice analysis: theory and application to travel demand. Transportation Studies (2018)
6. Benaglia, T., Chauveau, D., Hunter, D., Young, D.: mixtools: An R package for analyzing finite mixture models. Journal of Statistical Software (2009)
7. Blum, M., Nunes, M., Prangle, D., et al.: Comparative review of dimension reduction methods in approximate bayesian computation. Stat Sci **28**(2) (2013)
8. Breiman, L.: Random forests. Machine learning **45**(1), 5–32 (2001)
9. Brochu, E., Cora, V.M., De Freitas, N.: A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv preprint arXiv:1012.2599 (2010)
10. Chen, B., Castro, R., Krause, A.: Joint optimization and variable selection of high-dimensional gaussian processes. arXiv preprint arXiv:1206.6396 (2012)
11. Chu, Z., Cheng, L., Chen, H.: A review of activity-based travel demand modeling. CICTP 2012: Multimodal Transportation Systems pp. 48–59 (2012)
12. Davis, R., Lii, K., Politis, D.: Remarks on some nonparametric estimates of a density function. In: Selected Works of Murray Rosenblatt. Springer (2011)
13. Day, N.: Estimating the components of a mixture of normal components. Biometrika **56**(3), 463–474 (1969)
14. Gutmann, M.U., Corander, J.: Bayesian optimization for likelihood-free inference of simulator-based statistical models. J. of Machine Learning Research **17**(1) (2016)
15. Hutter, F., Hoos, H.H., Leyton-Brown, K.: Sequential model-based optimization for general algorithm configuration. In: International conference on learning and intelligent optimization. pp. 507–523. Springer (2011)
16. Izbicki, R., Lee, A.B., Pospisil, T.: ABC–CDE: Toward approximate bayesian computation with complex high-dimensional data and limited simulations. Journal of Computational and Graphical Statistics **28**(3), 481–492 (2019)
17. Järvenpää, M., Gutmann, M.U., Pleska, A., Vehtari, A., Marttinen, P., et al.: Efficient acquisition rules for model-based approximate bayesian computation. Bayesian Analysis **14**(2), 595–622 (2019)
18. Järvenpää, M., Gutmann, M.U., Vehtari, A., Marttinen, P., et al.: Gaussian process modelling in approximate bayesian computation to estimate horizontal gene transfer in bacteria. The Annals of Applied Statistics **12**(4), 2228–2251 (2018)
19. Jolliffe, I.T., Cadima, J.: Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **374**(2065) (2016)

20. Leclercq, F.: Bayesian optimization for likelihood-free cosmological inference. Physical Review D **98**(6) (2018)
21. Liaw, A., Wiener, M., et al.: Classification and regression by Random Forest. R news **2**(3), 18–22 (2002)
22. Lintusaari, J., Gutmann, M., Dutta, R., Kaski, S., Corander, J.: Fundamentals and recent developments in approximate bayesian computation. Syst Biol **66** (2017)
23. Lu, Y., Adnan, M., Basak, K., Pereira, F., Carrion, C., et al.: Simmobility mid-term simulator: A state of the art integrated agent based demand and supply model. In: 94th Annual Meeting of the Transportation Research Board (2015)
24. Marin, J.M., Pudlo, P., Robert, C.P., Ryder, R.J.: Approximate bayesian computational methods. Statistics and Computing **22**(6), 1167–1180 (2012)
25. Meinshausen, N.: Quantile regression forests. JMLR **7**, 983–999 (2006)
26. Močkus, J.: On bayesian methods for seeking the extremum. In: Optimization techniques IFIP technical conference. pp. 400–404. Springer (1975)
27. Nott, D., Fan, Y., Marshall, L., Sisson, S.: Approximate bayesian computation and bayes' linear analysis: toward high-dimensional ABC. Journal of Computational and Graphical Statistics **23**(1), 65–86 (2014)
28. Oh, S., Seshadri, R., Azevedo, C., Ben-Akiva, M.E.: Demand calibration of multimodal microscopic traffic simulation using weighted discrete SPSA. Transp Res Rec **2673**(5), 503–514 (2019)
29. Petrik, O., Adnan, M., Basak, K., Ben-Akiva, M.: Uncertainty analysis of an activity-based microsimulation model for Singapore. Future Gener Comp Sy (2018)
30. Price, L.F., Drovandi, C.C., Lee, A., Nott, D.J.: Bayesian synthetic likelihood. Journal of Computational and Graphical Statistics **27**(1), 1–11 (2018)
31. Qurashi, M., Maa, T., Chaniotakis, E., Antoniou, C.: PC-SPSA: Employing dimensionality reduction to limit SPSA noise in DTA model calibration. In: 2nd Symposium on Management of Future motorway and Urban Traffic Systems (2018)
32. Raynal, L., Marin, J., Pudlo, P., Ribatet, M., Robert, C., Estoup, A.: ABC random forests for bayesian parameter inference. Bioinformatics **35**(10) (2019)
33. Reynolds, D.A.: Gaussian mixture models. Encyclopedia of biometrics **741** (2009)
34. Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M.: Sensitivity analysis in practice: a guide to assessing scientific models, vol. 1. Wiley Online Library (2004)
35. Schultz, L., Sokolov, V.: Bayesian optimization for transportation simulators. Procedia computer science **130**, 973–978 (2018)
36. Sha, D., Ozbay, K., Ding, Y.: Applying bayesian optimization for calibration of transportation simulation models. Transportation Research Record (2020)
37. Sisson, S.A., Fan, Y., Beaumont, M.: Handbook of approximate Bayesian computation. CRC Press (2018)
38. Snoek, J., Larochelle, H., Adams, R.: Practical bayesian optimization of machine learning algorithms. In: NIPS. pp. 2951–2959 (2012)
39. Thomas, O., Pesonen, H., Sá-Leão, R., de Lencastre, H., Kaski, S., Corander, J.: Split-BOLFI for for misspecification-robust likelihood free inference in high dimensions. arXiv preprint arXiv:2002.09377 (2020)
40. Todorović, M., Gutmann, M., Corander, J., Rinke, P.: Bayesian inference of atomistic structure in functional materials. Npj computational materials **5**(1) (2019)
41. Wang, Z., Zoghi, M., Hutter, F., Matheson, D., De Freitas, N., et al.: Bayesian optimization in high dimensions via random embeddings. In: IJCAI (2013)
42. Wood, S.N.: Statistical inference for noisy nonlinear ecological dynamic systems. Nature **466**(7310), 1102–1104 (2010)