



# Black-Box Optimization for Buildings and Its Enhancement by Advanced Communication Infrastructure

Karel Macek<sup>a,c</sup>, Jiří Rojíček<sup>a</sup>, Georgios Kontes<sup>b</sup>, and Dimitrios V. Rovas<sup>d,b</sup>

<sup>a</sup> Honeywell Prague Laboratory

<sup>b</sup> Department of Production Engineering and Management, Technical University of Crete, Greece

<sup>c</sup> Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic

<sup>d</sup> Systems Integration Group, Fraunhofer Institute for Building Physics, Nurnberg, Germany

## KEYWORD

*Evolutionary algorithms*  
*Black box modeling*  
*Simplification*  
*Refining*  
*Building control*  
*HVAC*  
*Load shedding*  
*Communication infrastructure*

## ABSTRACT

*The solution of repeated fixed-horizon trajectory optimization problems of processes that are either too difficult or too complex to be described by physics-based models can pose formidable challenges. Very often, soft-computing methods - e.g. black-box modeling and evolutionary optimization - are used. These approaches are ineffective or even computationally intractable for searching high-dimensional parameter spaces. In this paper, a structured iterative process is described for addressing such problems: the starting point is a simple parameterization of the trajectory starting with a reduced number of parameters; after selection of values for these parameters so that this simpler problem is covered satisfactorily, a refinement procedure increases the number of parameters and the optimization is repeated. This continuous parameter refinement and optimization process can yield effective solutions after only a few iterations. To illustrate the applicability of the proposed approach we investigate the problem of dynamic optimization of the operation of HVAC (heating, ventilation, and air conditioning) systems, and illustrative simulation results are presented. Finally, the development of advanced communication and interoperability components is described, addressing the problem of how the proposed algorithm could be deployed in realistic contexts.*

## 1 Introduction

The use of advanced algorithms for scheduling and planning in real large-scale systems like buildings poses insurmountable challenges, related both to the large-scale but also the difficulty in establishing and using reasonably accurate first-principle models. In many cases – and the building case is one such example, – it is possible to consider a very detailed model of the system and – consequently – employ significant computational resources where an optimization problem would be solved in parallel. However, this approach is usually too expensive in terms of required configuration [QIN, S. *et al.*, 2003], [KORDON, A. 2007], [MAŘÍK, K. *et al.* 2011] and oftentimes the search process of a very large parameter space can fail to yield good performance that could lead to appreciable benefits and therefore would justify the significant efforts. Robustness and variability issues preclude such approaches from being applicable in realistic contexts. The use of

model-based approaches has attracted significant attention in the literature: model-predictive control approaches are currently under heavy investigation [OLDEWURTEL, F. *et al.*, 2012], [ZACEKOVA, E. *et al.* 2012], [HAVES, P. *et al.*, 2010] but the need for a devising a state-space model that can capture with accuracy all relevant dynamics often proves an insurmountable obstacle. The use of simulation-assisted approaches [KONTES, G. *et al.* 2012] where a detailed thermal simulation model is used, along with an optimization problem solved in a way that evaluations of the simulation model can be reduced, show some promise but nevertheless the complexity associated with the problem setup and managing the complex data flows, can be formidable. All such approaches, while interesting from a research perspective, are hard to transfer in commercial practice.

On the other hand, buildings account globally for approximately 40% of final energy use [ZACEKOVA, E. *et al.* 2012], and any steps towards optimization of building operation can



have significant economic and environmental impact. In developing such algorithms – if their applicability in practice is to be of relevance – they have to address three important issues: (i) such advanced control solutions for buildings has to be implementable in a resource-constrained hardware; (ii) the solution has to be deployable easily in a wide range of buildings; (iii) the solution has to still lead to more efficient operation compared to standard state of the art approaches. These goals seem to be contradictory: the limited hardware might lead to the need of more expert skills and more tailored installments, but these tailored installments might lead to complex deployment. The approach universally accepted in practice is the implementation of a set of simple decision rules, e.g. [YU, Z. *et al.* 2010], which typically do not perform optimization of any kind. What is equally egregious is that the complexity of integration, not only limits the intelligence that can be implemented, but oftentimes errors in the rules or incorrect decisions can lead to erratic or even poor energetic performance.

The complexity associated with programming the hosting platforms (usually the Programmable Logic Controllers in the building) creates a number of practical issues that hamper practical testing of such approaches. Light-weight software platforms, capable of managing complex information flows and providing hosting capabilities for such energy efficiency services are currently in development [ROJÍČEK, J. *et al.* 2013] and can help streamline the deployment and testing process.

In the present paper an approach is presented that help reconcile some of the aforementioned issues in a practical and real-world applicable fashion. To that effect, we introduce an algorithm to optimize costs related to the operation of a building using only a limited set of available data. In the remaining of the paper we discuss first, in Section 2, the solutions available for solving a dynamic decision problem and introduce important building blocks. Using these blocks, in Section 3, the proposed algorithm is presented. In Section 4 a simulation-based analysis of the algorithm is performed. Then in Section 5, we discuss the communication and configuration requirements for implementation of the proposed algorithm in realistic contexts.

## 2 Black-Box Modeling and Optimization in Buildings

With respect to the goals mentioned in Section 1, namely obstacles related to the use of first-principle modeling approaches, a black-box modeling and optimization seems more fitting for our study. In this section we also provide basic information the addressed domain: the pre-cooling control of air-conditioning systems.

Symbol	Explanation
$x \in X$	Input
$y \in Y$	Output
$\mu(x)$	Mean value of $y x$
$\sigma(x)$	Standard deviation of $y x$
$s \in S$	State
$u \in A$	Action
$\epsilon$	Noise, i.e. $y - \mu(x)$
$u^*$	Optimal action
$\alpha$	Risk factor
$h \in H$	Set of simplified states
$C \subset A$	Restricted set of actions
$b \in B$	Set of simplified actions
$p$	Mapping simplifying actions
$r$	Mapping simplifying states
$\mathbf{N}$	Set of natural numbers
$\mathbf{R}$	Set of real numbers
$\mathbf{R}_+$	Set of non-negative real numbers
$d$	Data $d = (x(j), y(j), a(j))_j$
$j$	Time of decision making
$T_{i,t}$	Zone air temperature setpoint at $t$ in the $i$ th zone

Tab. 1: List of symbols

### 2.1 Black-Box Modeling

There exists a variety of methods that can be considered as black-box models, e.g. neural networks [ROJAS, R. 2006], Gaussian processes [RASMUSSEN, C. *et al.* 2006], or local regression [WASSERMAN, L. 2006]. These methods typically focus on modeling the dependence between inputs (or regressors)  $x$  and outputs  $y$  with zero-mean noise  $\epsilon$  and finite variance  $\sigma^2$ .

$$y = \mu(x) + \epsilon \quad (1)$$

Note the variance can be input-dependent, i.e.  $\sigma = \sigma(x)$ . This dependency can be identified using past data  $d = (x(j), y(j))_{j=1..m}$ . For our purposes, we extend the concept to the conditions  $s$  and actions  $a$ .

$$y = \mu(s, a) + \epsilon \quad (2)$$

Therefore, we re-define the data set as  $d = (s(j), a(j), y(j))_{j=1\dots m}$ . From the variety of possible approaches to black-box modeling, we decided for the local regression which is described in detail in [WASSERMAN, L. 2006].

## 2.2 Black-Box Optimization

The black-box optimization consists in solving the following problem:

$$u^* = \arg \min_{u \in U} f(u) \quad (3)$$

where  $s_{now}$  stands for the actual state.

Also in this case a wide range of methods is available: Differential Evolution [STORN, R. *et al.* 1997], Particle Swarm Optimization [KENNEDY, J. *et al.* 1995] or Covariance Matrix Adaptation [HANSEN, N. *et al.* 1996]. In this paper, we decided for Covariance Matrix Adaptation.

We intend to use the model (2) for the black-box optimisation (4). For this purpose we will fix the state variable  $s = s_{now}$  to the actual value. For dealing with uncertainty noise, we use

$$f(a) = \mu(a, s_{now}) + \alpha \sigma(a, s_{now}) \quad (4)$$

where  $\alpha$  is a real parameter. At the beginning,  $\alpha$  can be very high which leads to exploration of the input space.

## 2.3 Case Study: Optimal Pre-Cooling

Let us consider the particular application of modeling (2) and optimization (4) in buildings. The problem we intend to solve is that of load-shedding through the utilization of pre-cooling. Pre-cooling is a strategy where the operation of an air conditioner is performed during off-peak periods (typically nights or early in the morning) so that the spaces can be preconditioned before the actual need arises. Operating in such a mode, leads to lowering indoor-air temperature and reduction of the temperature in surrounding thermal storage elements, therefore requiring less (in some cases, no) effort during peak electricity demand periods. In that sense the problem we are addressing in the present work is the determination through the solution of the optimization problem in (3) of pre-cooling strategies to be applied on a 24-hour window into the future that will lead to cost reduction while maintaining thermal comfort within acceptable levels. We assume this decision regarding next day operation occurs at midnight and then the strategy devised is used till next midnight when a new solution of the optimization problem is performed and a new strategy is devised for the following day.

We formulate the problem (3) where the actions  $a$  are the internal temperature set-points over the next  $t_{max}$  time intervals in  $n$  considered zones, i.e.  $T_{t,i}$ . We consider some comfort limits for those temperatures which can be zone and time<sup>1</sup> specific  $T_{t,i}^{min} \leq T_{t,i} \leq T_{t,i}^{max}$  where  $t = 1, \dots, t_{max}$  and  $i = 1, \dots, n$ . These constraints define the set  $A$  of allowable actions. The state  $s$  involves all relevant information which is important for the decision making, namely:

- Information about the actual state of the building. This can be based on available measurements as well as on more sophisticated filtration.
- Information about the anticipated influences on the building. This involves the weather forecast (temperature, humidity, solar radiation, wind speed and orientation), possibly also the occupancy forecast and profile of dynamic electricity prices.

Finally, the output  $y$  is the overall cost for given  $t_{max}$  time instants. Each data record will relate to one day and contain  $a$ ,  $s$ , and  $y$ .

Now, let us analyze the typical complexity of this problem. Let us consider the horizon 24 hours with 15 minutes sampling, 10 zones, 5 external characteristics, 5 variables describing the actual building state. The regression model, will have  $24 \cdot 4 \cdot (10+5) + 5 = 1445$  input variables. In order to obtain sufficient level of precision, a large training set and non-trivial computing time would be needed. The number of action variables equals in this case  $24 \cdot 4 \cdot 10 = 960$  and makes the optimization problem very complex. Existing efforts of black-box modeling and optimization in the context of building control seem to be implicitly aware of these complexity limitations since they are mostly focused on the small scale cases only: see e.g. [GANCO, M. *et al.* 2012], [DORVEL, K. *et al.* 2012], or [HU, M. *et al.* 2012]. There are already attempts on making the pre-cooling problem more computationally tractable, e.g. by means of Reinforcement Learning [HENZE, G. *et al.* 2006]. However, also in case of application of Reinforcement Learning the problem remains very complex. A realistic use of the black-box modeling in the building can be found in [MACEK, K. *et al.* 2012] where two control strategies are compared over a longer time

<sup>1</sup> Note that  $t$  is used as a way of indexing during a single day, e.g.  $t = 10$  denotes the time interval between 10 and 11 pm. In contrast to  $k$  which is used for the iteration of the decision problem (4) and in this case is used for indexing of whole days.

horizon with respect to the external changing conditions. This led us also to the idea to use the evaluation of the past high-level decisions for making new decisions which is subject of this publication; its first version was published in [MACEK, K. *et al.* 2013].

### 3 Black-Box Approach with Refining

The core idea presented in the paper consists in starting the decision making with very limited information and simple models which can be made more complex and accurate with increasing size of data available. First, we define the concept of simplifications for the decision problems. Afterwards, the basic version of the algorithm is presented. Consequently, we introduce the refining, the ways of testing it, and the ways to implement it.

#### 3.1 Concept of Simplification

The goal of simplification is to make the decision problem less complex and computationally feasible. The price paid for the simplification is the information loss. The simplification is used for both actions  $a$  as well as for the states  $s$ . Both cases differ in the way how the simplification is carried out.

In case of states  $s$ , the simplification is a mapping to a reduced-dimensional space, i.e.  $r: S \mapsto H$ , which can be implemented using e.g. Principle Component Analysis [WOLD, S. 1987] or Feature Selection [GUYON, I. 2003]. By means of the  $r$  any point from  $S$  can be mapped to its simplified version.

In case of actions, we use a slightly different approach. Instead of the set of high dimensional vectors  $A$ , we use a subset  $C \subset A$ . This practically means that we restrict our decisions and some of them cannot be made with the given simplification. During the model identification, the records where the actions are not from  $C$  are omitted. Even though  $C$  can be a small subset of  $A$ , it is still high dimensional. Therefore we introduce a bijective mapping  $p$  maps  $C$  to  $B$  which is of a lower dimension than  $A$ . An example of simplification of the actions is the case where most components are fixed and some of them are of the same value. In this case  $B$  is a set of such values and is one-dimensional. Based on the previous paragraph, we define that a simplification as a 5-tuple  $z = (r, H, C, B, p)$ . The set of all simplifications will be denoted as  $Z$ .

The application of a simplification to data means:

1. Removing all data where  $a(j)$  is not in  $C$ .
2. Mapping each record's state to  $h(j)$  using  $r$ .
3. Mapping each record's action to  $b(j)$  using  $p$ .
4. Saving the record  $(h(j), b(j), y(j))$  into the new data set  $d'$ .

The simplification is illustrated in Figure 1. It is obvious, that some simplifications can be more suitable than others. If we fix the some components of  $a$  which have the most significant impact on the output, the simplification will lead us to worse solutions than if we fix the less important components and optimize the more important ones. Similar situation is in case of the states: simplifying the data to the first component of the PCA analysis can be assumed to be more suitable than use of the very last component.

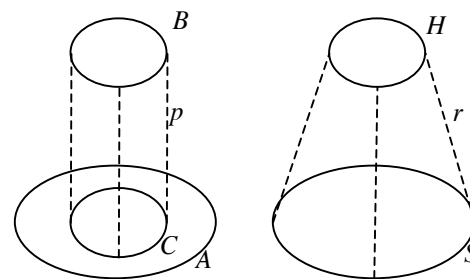


Fig. 1. Simplification of states and actions. The ovals represent the sets while the dashed lines the mappings  $p$  and  $r$ .

Let us describe the simplification in the considered case study. Assume that we have a total of  $n$  zones, and our decision problem is to identify the temperature setpoint trajectories for each of the zones:  $T_{i,t}, i \in \{1, \dots, n\}$ . These can be quite general functions but to much the application in reality can be approximated as piecewise constant functions. In reality we are not allowed to change the setpoints as often as we would like and we should keep the values constant over finite time intervals. We therefore introduce a partition of the time interval we are interested in, and subdivide the 24h interval into equally spaced hourly intervals. The assumption here is that we plan to be taking an action every hour, altering the setpoint appropriately; this is a quite reasonable approximation given the slow dynamics of the problem we are trying to address.

We define the set of actions  $A$  as the setpoint variations on each of the zones, using this discretized approximation of setpoint variations

on each of the zones. The actions  $A$  will be restricted to the following subset  $C$ :

$$T_{t,i} = \begin{cases} 22^\circ\text{C}, & \forall t \in \{9, \dots, 22\} \\ 30^\circ\text{C}, & \forall t \in \{23, 24\}, i \\ b, & \forall t \in \{1, \dots, 8\} \\ & = 1 \dots n \end{cases}$$

where  $15^\circ\text{C} \leq b \leq 30^\circ\text{C}$ .

Using this simplification it is easy to see that the set  $B = [15, 30]$  is our search space and a point there represents one feasible solution to our problem. In case of  $S$  we will use the first PCA component from the data containing weather forecasts and actual states in the building. As we can see, we face a two dimensional problem, now.

### 3.2 Basic Algorithm

The basic approach can be summarized in the following steps:

1. Define the simplification  $z$  from  $Z$ .
2. Define the data set as empty, i.e.  $d = \emptyset$ .
3. Set  $j = 1$
4. Generate a random action  $a(1)$  from  $A_1$ , store it the related state  $s(1)$ .
5. Apply the action  $a(j)$  to the system and observe the costs  $y(j)$ .
6. Extend the data set  $d$  by new  $(s(j), a(j), y(j))$ .
7. Obtain and store new state  $s(j+1)$ .
8. Optimize new action  $a(j+1)$  for given  $z$ ,  $d$ , and  $s(j+1)$ . This involves:
  - (a) Applying  $z$  to  $d$  and produce  $d'$ .
  - (b) Identification of the model  $\mu(h, b)$ ,  $\sigma(h, b)$  of the loss function using the black-box modelling from data  $d'$ .
  - (c) Formulation of  $f(b)$  based on  $\mu$ ,  $\sigma$ , and  $r(s(j+1))$ .
  - (d) Solution of  $b^* = \operatorname{argmin} f(b)$
  - (e) Putting back from the simplified world, i.e.  $a(j+1) = p^{-1}(b^*)$
9. Set  $j = j + 1$  and go to 5.

Let us see the function  $\mu$  for our case study when  $j = 20$  in Figure 2. It seems that for lower value of  $h$ , the pre-cooling is more desired while in case of high values, pre-cooling is not economical. Note that for values  $h < -5$ , there are only two data points available, therefore the model shall be considered as approximate only.

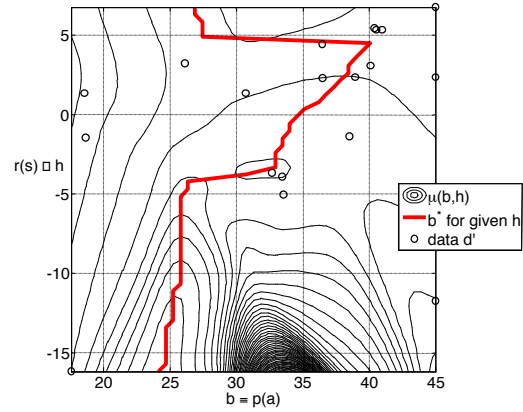


Fig. 2: Black-box model for the proposed simplification in the decision problem and optimal simplified actions  $b^*$  for given simplified conditions  $h$ .

### 3.3 Concept of Refining

The simplification we propose allows us to make a decision problem significantly easier. In an extreme case we can use one-dimensional  $B$  and one-dimensional  $H$  which is computational feasible and only few data records are required for the model identification. The only issue is the information loss and limited decision space. After more data are available, it more complex models are worth considering.

For this purpose we introduce the concept of refining of simplifications. Let us have two simplifications  $z_1 = (r_1, H_1, C_1, B_1, p_1)$  and  $z_2 = (r_2, H_2, C_2, B_2, p_2)$ . We say that  $z_2$  refines  $z_1$  if and only if:

- $C_1$  is a subset of  $C_2$ .
- There is a mapping  $g: H_2 \rightarrow H_1$  such that  $r_1(s) = g(r_2(s))$  for all  $s$ .

The concept of simplification is shown in Figure 3.

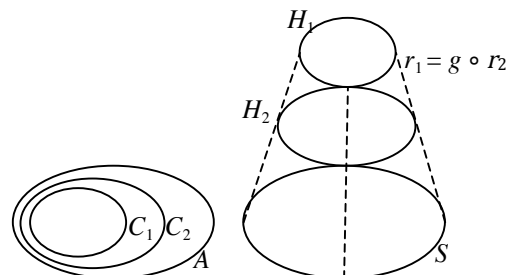


Fig. 3. Refining the simplification  $z_1$  to more informative simplification  $z_2$ .

In our case study, we can consider the following refining: Let  $C_1$  is as introduced in 1.1. The  $C_2$  will be defined by the same constrains as  $C_1$ , only the last one will be replaced by:

- at time instants the  $t = 1, 2, \dots, 4$  the temperature setpoint is the same, i.e.  $T_{t,i} = T_{k,i}$  for all  $t, k = 1, \dots, 4$

- at time instants the  $t = 5, 6 \dots 8$  the temperature setpoint is the same, i.e.  $T_{t,i} = T_{k,i}$  for all  $t, k = 5, \dots 8$

We can insight that this we can refine this simplification again: from quadruples to pairs and then from pairs to individual components.

### 3.4 Tests on Refining

Before refining the actual simplification, it has to be assured that the procedure achieves for the given simplification the best possible results. For this purpose we propose the application of the following test. However, other tests can be considered, too. Let  $\rho \in \mathbf{R}_+$ , let  $\theta \in \mathbf{N}$ . Let  $w(j)$  be defined as follows:

$$w(j) = \begin{cases} 0 & \text{if } j = 1 \\ w(j-1) + 1 & \text{if } \|b^* - b_{prev}\| < \rho \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where  $b_{prev}$  is the optimal solution using the actual state  $s(j)$  and data without the very last record ( $s(j-1), a(j-1), y(j-1)$ ).

The refining is carried out after  $w(j) > \theta$ . It means when the new data record did not provide significant information to the problem. The value of  $\theta$  will be discussed in numerical examples in Section 7.

### 3.5 Implementation of Refining

There are typically many ways how to carry out the refining. In case of states  $S$ , it is possible to involve more components from the Feature Selection of Principal Component Analysis.

In case of actions  $A$ , the refining can relax a constraint of the set  $C$ . Since it can be described by a set of equation and inequation, the refining can consist in omitting some of them. If we consider that  $C_1$  is a set of actions where most components are of fixed values and the rest has one value which can be within an interval. A simple refining can be done by dividing these variable components into two groups and considering that each group has its own variable value. Section 4 contains an example of that.

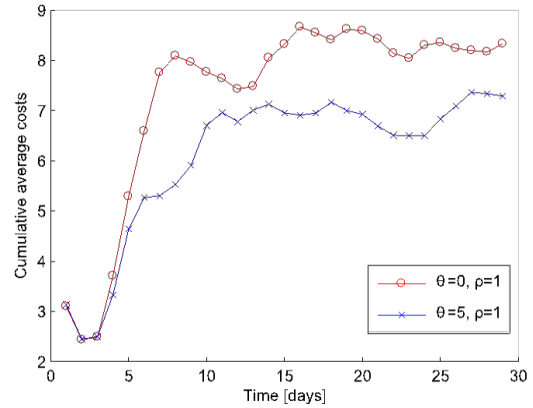


Fig. 4. Comparison of fast and slow refining. It is visible that the slower refining leads to better results than faster refining that is closer to the approach without simplifications.

## 4 Results

### 4.1 Details on the Case Study

The case study has been introduced step-by step in the previous sections. In order to demonstrate the results numerically, we provide basic information about the set-up since detailed description is beyond the scope of this work. At the end of this section, achieved results are discussed.

First, a first-principle simulation model has been adopted for a single zone building with one chiller. We adopted a version of [STRELEC, M. *et al.* 2012] where the thermal capacity of the zone is influenced by ambient air temperature, heat gains from the occupants (people in the zone), internal thermal inertia given by gains and the chiller itself. The parameters have been determined based on experience of modeling of large-scale single zone buildings. The weather profile has been used from a real building, slightly shifted so it had values between 20 and 35 °C.

### 4.2 Specification of the Algorithm

The adopted black-box model was based on local polynomial regression with Gaussian kernel with covariance matrix  $\Sigma_{i,i} = 1$  and  $\Sigma_{i,j} = 0$  if  $i \neq j$  and degree 2 (quadratic regression). More details can be found in [WASSERMAN, L. 2006]. The considered model had single output, namely the costs  $y$  and several regressors: first was related to the state  $s$  while the others represented action  $a$ . The first regressor was the first PCA [WOLD, S. 1987] component from data containing: (i) weather

profile from the last day and (ii) internal temperatures (zone air, building construction). As a procedure for optimization of  $b = p(a)$ , the covariance matrix adaptation [HANSEN, N. *et al.*1996] was selected and the limit of iterations was 200 for each optimization. We worked with 30 days and carried out 2 experiments, one for  $\theta = 0$ , another for  $\theta = 5$ . In both cases,  $\rho = 1$  and  $\alpha = 0$ .

### 4.3 Results

Figure 4 illustrates the results. The first experiment with the setting  $\theta = 0$  leads to fast refining and the problem has comparable complexity as if the 8-dimensional problem would be addressed directly. The other experiment with  $\theta = 5$  leads to slower refining. It can be observed that better results, i.e. lower costs, are obtained when the dimensionality grows slowly with increasing information in the data.

## 5 Conclusions

### 5.1 Summary of Achievements

In this article we have analysed the state of the art in black-box approaches in the optimal decision making for building system. In particular we focused on the task of pre-cooling which is of significant importance both with respect to economic and environmental objectives.

Based on the state of the art analysis, our experience from real buildings as well as knowledge of business needs, we have proposed a novel method of refining which is able to start with a very small data sets and decisions, which are based on models over simplified data set. With increasing size of data available, the applied simplifications can be refined and lead to better decisions.

We have proposed also a simple experiment based on a simulated building. This experiment confirmed our hypothesis that optimisation with incremental refining leads to better results than direct optimisation of a high dimensional problem.

### 5.2 Advantages and disadvantages of the proposed method

Let us start with the advantages of the proposed solution. First, the method is very general, not limited to the pre-cooling only. The method works incrementally in the top-down way. Even though it is based on the black-box approaches,

the user can use the prior knowledge in the definition of the initial simplification and consequent refining. Another way how the method can use the prior knowledge is to incorporate some artificial data into the used data set. These data can be generated e.g. from the established laws of physics.

On the other hand, the method has some drawbacks, too. In case of an inappropriate simplification, the black-box models will be of very non-linear shapes or of large noise. Next, the approach is a kind of supervisory decision making and has to rely on the lower level control. Some issues can be assumed also in case of the proposed testing for refining. New data might lead to no improvements just because these data are too closed to already measured data. After a while, new data can occur and the given simplification can learn something new.

### 5.3 Solution within the BaaS Architecture

Apart from the requirements for proper definition of the simplifications and availability of lower-level control logic, wide adoption of the proposed methodology by stakeholders managing real buildings is hindered by the complexity associated with the problem of setting-up and managing the complex data flows, while existing BEMS are unable to support the programming and deployment of such model-based approaches.

In an effort to facilitate the transition to advanced supervisory control solutions in real buildings, the utilization of BaaS “smart platform” allows smooth integration, testing and deployment of such approaches. Here, a three layer architecture has been implemented (Figure 5), consisting of: the Data Layer, serving static and dynamic data through a set of data sources or repositories (a data warehouse, a Building Information Management (BIM) server, a Building Management System (BMS) and external services); the APO Service Layer, providing support for hosting the reasoning and analytics services; and the Communication Logic layer acting as an abstraction layer to facilitate communication between the Data Layer and the APO Services Layer. These functionally disjoint layers operate independently, communicating through the use of properly-defined (software) interfaces. For more information on the overall BaaS architecture, please refer to [VALMASEDA, C, *et al.* 2013].

Within this platform, the APO Service Layer is responsible for hosting the components of the proposed methodology, i.e. the optimization algorithm and the regression model utilized, as different *modules* subscribed to the system. Initially, all necessary parameters for the deployment of the modules (e.g. the simplifications, the inputs to the PCA algorithm, etc.) are configured through a GUI tool, and the modules are registered to the system. Subsequently, the building manager defines the execution interval (e.g. every night at 00:00) of the overall optimization methodology, while the data management and access services are provided in a transparent way from the platform. To that direction, two sets of data types are required: *static data*, defining the mapping between the requested module inputs/outputs and the available data sources; and *dynamic data*, containing all the necessary time-dependent data. For the first category, the description of the building is available in IFC4 format [ISO 16739, 2013] through a BIM repository, and using a set of queries to the BIM automatic configuration of the modules is possible, while for the second category, the Data Warehouse and external sources provide the necessary historical and forecasted data

respectively – see [VALMASEDA, C. *et al.* 2013] for an example of automatic module configuration using a BIM repository. In addition to the semi-automatic configuration of the overall process, the BaaS platform offers a plethora of pre-configured FDD analytics, ensuring the necessary data and service integrity for the successful deployment of the proposed solution. Such functionality is essential, since the accuracy of the regression model depends on the quality of the historical and forecasted data, while the efficiency of the final solution can be degraded by system faults or energy-prone user behavior – see [ROJÍČEK, J. *et al.* 2013] for an example of a BaaS FDD module. To conclude, the utilization of BaaS platform allows smooth integration of the proposed methodology in real buildings, thus enabling its adoptability by the market, while the selected modular architecture allows rapid deployment, integration and evaluation of different modules for the same task (e.g. evaluating the performance of different regressors), thus providing increased expandability and offering a holistic deployment and evaluation suite.

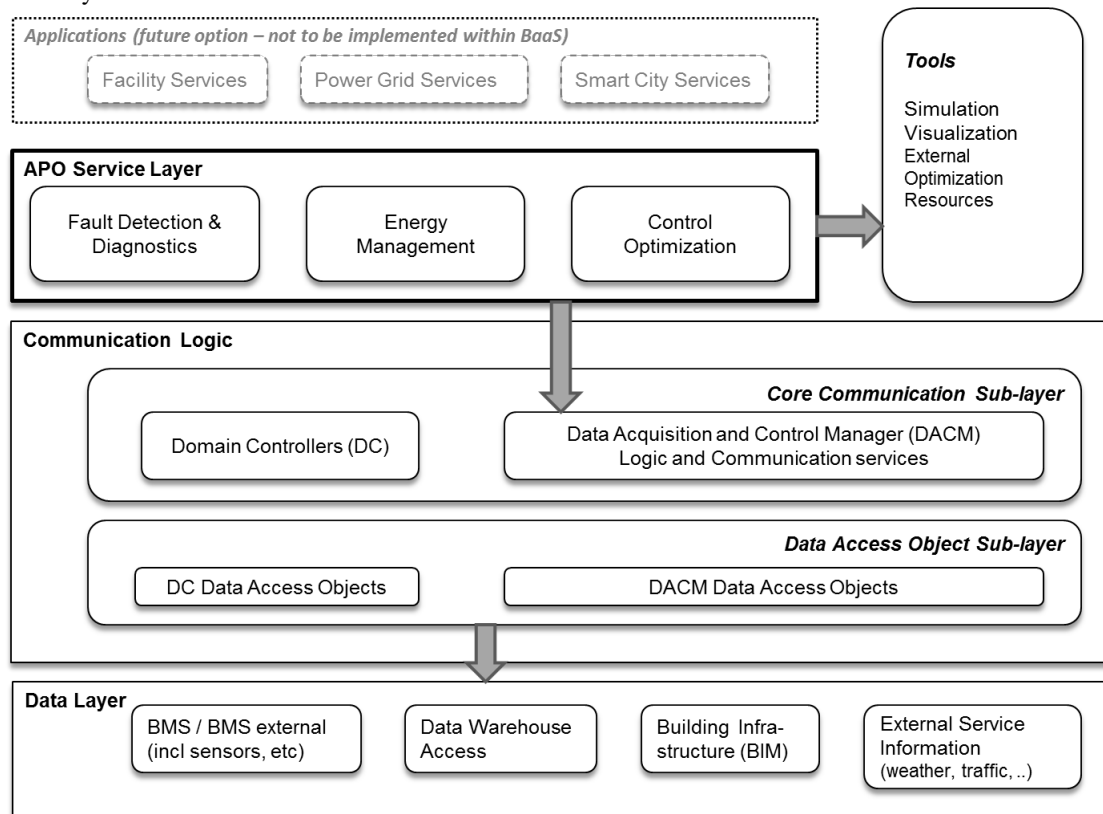


Fig. 5. BaaS high-level architecture



## 5.4 Other Future Work

The proposed method of simplifications and refining opens a wide range of next research challenges.

The first group is related to minor incremental activities. The method shall be tested also in the real operation or at least in more sophisticated simulation models [DOE]. For the testing other black-box models and other black-box optimization algorithms as well as other refining and simplifications can be considered. Another possibility is to test the proposed approach in a completely different field. Next, the approach shall challenge other state-of-the-art approaches, e.g. the Model Predictive Control or Reinforcement Learning.

The other group consists in more complex topics. The theoretical analysis should examine the complexity aspects, accuracy of the simplified decision making, as well as the convergence to the original problem. Another topic is to consider the addressed problem as a dynamic one instead of an iterated static problem. The dynamics can be considered at different levels (hours, days ...). Finally, is worth considering to apply the approach in a distributed way where each agent would be represented e.g. by a single zone and actions of one agent could be states for some other agents.

## 6 Acknowledgment

The research leading to this work has been supported by the European Commission FP7-2011-6, ICT Systems for Energy Efficiency under contract #288409 (BaaS – Building as a Service) and Grant Agency of the Czech Republic GACR 13-13502S.

## 7 Vitae



Karel Macek received a BS in Applied Mathematics in 2004 from the Czech Technical University in Prague (Czech Republic) and a MS in Theoretical Computer Science from the Charles University in Prague (Czech Republic). From 2006 to 2007, he worked as a consultant in performance improvement of business processes

for PricewaterhouseCoopers. Since March 2008, he has been a Research Engineer at the Honeywell Prague Laboratory (Czech Republic). His research interest is in the application of soft-computing methods in energy systems and building automation.



Jiri Rojicek received his Ph.D. degree in Mathematical Engineering from the Czech Technical University in Prague in 1999. His research interest is building optimization, especially focused on fault detection and diagnostics, performance monitoring of HVAC equipment and building energy management. He is an author and co-author of 20+ papers and 4 patents. Since 1999 he is with Honeywell Prague Laboratory, currently as Principal Research Engineer.



Giorgos Kontes received the BSc degree in Electronics and Computer Engineering from the Department of Electronic and Computer Engineering of Technical University of Crete in 2007 and the MSc degree in Production Systems from the Department of Production Engineering and Management of Technical University of Crete in 2010. Since 2010, he is a Phd student in the Department of Production Engineering and Management of Technical University of Crete. His research interests include the application of Reinforcement Learning and Stochastic Optimisation methods for Intelligent Design of Building Energy Management Systems.



Dimitrios Rovas received a B.Sc. in Mechanical Engineering from the National Technical University of Athens (Greece) and a Ph.D in 2003 in Mechanical Engineering from the Massachusetts Institute of Technology. He has worked as an Assistant Professor at the University of Illinois at Urbana-Champaign

performing research in the area of Computational Science and Engineering. He then moved to the faculty of engineering at the Technical University of Crete, where currently he is an adjunct faculty member. Since 2013, he

is a member of the Systems Integration Group of the Fraunhofer Institute for Building Physics. His research interests are in the area of simulation and optimization of energy systems.



## 8 References

- [DORVEL, K. *et al.* 2012] K. DORVEL, S. MEDVED, Multi-objective optimization of a building free cooling system, based on weather prediction, *Energy and Buildings*, 52 (2012), 99-106
- [GANCO, M. *et al.* 2012] M. J. GANCO, R. ALCALA, F.HERRENA, A multi-objective evolutionary algorithm for an effective tuning of fuzzy logic controllers in heating, ventilating and air conditioning systems, *Applied Intelligence*, 36 (2) (2012), 330-347
- [GUYON, I. *et al.* 2003] GUYON, I., ELISEEFF, A., An introduction to variable and feature selection, *The Journal of Machine Learning Research*, 3 (2003) 1157-1182
- [HANSEN, N. *et al.*1996] N. HANSEN, A. OSTERMEIER, Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. in: *Proceedings of IEEE International Conference on Evolutionary Computation*, Nagoya, Japan, 1996, pp. 312-317
- [HAVES, P. *et al.*, 2010] P. HAVES, B. HENCEY, F. BORELLI, J. ELLIOT, Y. MA, B. COFFEY, S. BENGEEA, M. WETTER, Model predictive control of HVAC systems: implementation and testing at the University of California, Merced, technical report, Lawrence Berkeley National Laboratory, Berkeley, California, 2010.
- [HENZE, G. *et al.* 2006] G. P. HENZE, S. LIU, Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory: Part 1. theoretical foundation, *Energy and Buildings*, 38 (2) (2006), 142-147
- [HU, M. *et al.* 2012] M. HU, J. D. JEFFERY, T.WU, Peak load shifting control using a mimetic algorithm, *European Journal on Operational Research*, 217 (1) (2012), 185-197
- [ISO 16739, 2013] ISO 16739 (2013). Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries. The International Organization for Standardization
- [KENNEDY, J. *et al.* 1995] J. KENNEDY, R. C. EBERHART, Particle swarm optimization, in: *IEEE International Conference on Neural Networks*, Piscataway, New Jersey, 1995, pp. 1942-1948
- [KONTES, G *et al.* 2012] G. KONTES, G. GIANNAKIS, E. KOSMATOPOULOS, D.V. ROVAS, Adaptive fine-tuning of building energy management systems using co-simulation, in *Proceedings: 2012 IEEE International Conference on Control Applications*, Dubrovnik, Croatia, 2012, pp. 1664-1669.
- [KORDON, A. 2007] A. K. KORDON, *Applied computational intelligence – how to create value*, Springer, Berlin Heidelberg, 2007.
- [MACEK, K. *et al.* 2012] K. MACEK, K. MAŘÍK, A methodology for quantitative comparison of control solutions and its application to HVAC (heating, ventilation and air conditioning) systems, *Energy*, 44 (1) (2012) 117-125
- [MACEK, K. *et al.* 2013] K. MACEK, V. BIČÍK, J. ROJÍČEK, Trajectory optimization under changing conditions through evolutionary approach and black-box models with refining, *Advances in Intelligent Systems and Computing* 217 (2013) 267-274
- [MAŘÍK, K. *et al.* 2011] K. MAŘÍK, J. ROJÍČEK, P. STLUKA, J. VASS, Advanced HVAC Control: Theory vs. Reality, in: *Preprints of the 18th IVAC World Congress*, Milano, Italy, 2011, pp. 3108-3113
- [OLDEWURTEL, F. *et al.*, 2012] F. OLDEWURTEL, A. PARISIO, C. N. JONES, D. GYALISTRAS, M. GWERDER, V. STAUCH, B. LEHMANN, M. MORARI, Use of model predictive control and weather forecasts for energy efficient building climate control, *Energy and Buildings*, 45 (2012), 15-27
- [QIN, S. *et al.*, 2003] S.J. QIN, T.A. BADGWELL, A survey of industrial model predictive control technology *Control Engineering Practice*, 11 (2003) 733–764
- [RASMUSSEN, C. *et al.* 2006] C. RASMUSSEN, C. K. I. WILLIAMS, *Gaussian processes for machine learning*, The MIT Press, Cambridge, Massachusetts, 2006
- [ROJAS, R. 2006] R. ROJAS, *Neural networks – a systematic introduction*, Springer, Berlin,

- 1996
- [ROJÍČEK, J. *et. al* 2013] J. ROJÍČEK, R. FIŠERA, G.D. KONTES, G.I. GIANNAKIS, D.V. ROVAS, Functional and technological definition of BIM-aware services to Assess, Predict and Optimize energy performance of buildings. In Proceedings 2<sup>nd</sup> Central European Symposium on Building Physics, Vienna, Austria, September 9-11, 2013. (To Appear)
- [STORN, R. *et al.*1997] R. STORN, K. PRICE, Differential evolution – a simple and efficient heuristic for global optimization, *Journal of Global Optimization*, 11 (1997), 341-359
- [STŘELEČEK, M. *et. al.* 2012] M. STŘELEČEK, K. MACEK, A. ABATE, Modeling and simulation of a microgrid as a stochastic hybrid system, in: *Proceedings of the IEEE PES Innovative Smart Grid Technologies*, Berlin, Germany, 2012, pp. 1-9.
- [VALMASEDA, C. *et al.* 2013] C. VALMASEDA, K. KATSIGARAKIS, M.A. GARCIA FUENTES, J. L. HERNANDEZ GARCIA, G.D. KONTES, D.V. ROVAS, An event-driven SOA-based platform for energy-efficiency applications in buildings, *CIBW78 2013*, Beijing, China (To appear).
- [WASSERMAN, L. 2006] L. WASSERMAN, *All of nonparametric statistics*, Springer, New York, 2006
- [WOLD, S. 1987] S. WOLD, Principal component analysis, *Chemometrics and Intelligent Laboratory Systems*, 2 (1) (1997), 37-52
- [ZACEKOVA, E. *et al.* 2012] E. ZACEKOVA, L. FERKL, Building modeling and control using multi-step ahead error minimization, in: *Proceedings of Mediterranean Conference Control & Automation*, Barcelona, Spain, 2012, pp. 421-426.