Automation in Construction 14 (2005) 551–568

AUTOMATION IN CONSTRUCTION

www.elsevier.com/locate/autcon

Analytic target cascading in simulation-based building design

R. Choudhary[a,*], A. Malkawi[b], P.Y. Papalambros[c]

[a]College of Architecture, Georgia Institute of Technology, Atlanta, GA 30332-0155, USA
[b]Department of Architecture, University of Pennsylvania, Philadelphia, PA 19104-6311, USA
[c]Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109-2250, USA

Accepted 28 November 2004

Abstract

This article presents a rigorous simulation-based optimization framework that enables concurrent and consistent decision-making in building design. Analytical Target Cascading (ATC), a multi-level engineering design optimization framework, is extended to thermal and HVAC design in buildings. The framework facilitates computational decision support for meeting building performance goals, allows autonomy of specialized design tasks with timely and efficient use of analysis tools, and preserves dependencies between possibly competing building performance goals. A pilot application demonstrates how ATC functions in the context of building design. Relevance and benefits of this hierarchical optimization approach to multi-criteria building performance problems are also discussed.
Published by Elsevier B.V.

Keywords: Simulation-based design; Optimization; Building performance; Analytic Target Cascading

1. Introduction and background

Building simulation's central concern is design performance. The field of building simulation is dedicated to developing analytic tools for modeling and computing performance in design; it has significantly influenced design practice and the way that computational analytic tools have been used to examine a design's performance. Hence, the iterative process of making design decisions, computing their effects on design performances, and evaluating and comparing the results with previous design decisions is well integrated in the use of building simulation tools. Design optimization can formalize and improve this process.

Optimization techniques have been used in architecture primarily for solving problems of space layout, structural design, and building performance. Space layout optimization is concerned with finding feasible topology and dimensions of interrelated objects that meet all design requirements and maximize design preferences [16,17]. Structural optimization involves configuration of structural elements and whole con-

* Corresponding author. Tel.: +1 404 385 2913; fax: +1 404 894 0572.
E-mail address: ruch@coa.gatech.edu (R. Choudhary).

0926-5805/$ - see front matter. Published by Elsevier B.V.
doi:10.1016/j.autcon.2004.11.004

structions, their geometrical and mechanical features, and the properties of their materials for optimizing mechanical, economic and/or aesthetic criteria [2]. Building performance optimization involves positioning and sizing architectural elements, selecting materials, and determining appropriate control settings for maximizing acoustic, thermal or lighting performance criteria in a given context. Performance optimization is different from solving structural design and space allocation problems. In addition to geometry, topology, and materials, it is significantly influenced by schedules of control and operation of the design. Similar to structural design, model functions in building performance optimization problems are mostly simulation-based, often involving complex finite element analysis.

Gero et al. [9] and Wilson and Templeman [36] were among the first to present building performance problems within the design optimization framework. Wilson and Templeman [36] derived design decisions that minimize initial and operational costs of an office building by posing the problem as a constrained nonlinear optimization problem. Gero et al. [9] included energy efficiency in the context of other performance criteria and proposed a design optimization model for generating Pareto sets to understand trade-offs between multiple performance criteria.

In general, problems involving architectural design elements are often ill defined. Too often, design elements are selected by simultaneously considering numerous quantifiable as well as non-quantifiable criteria. In addition, the nature of the problems and use of complex simulations to evaluate functions often yield undesirable properties in the optimization model. For these reasons, the use of numerical methods to solve building performance problems has been restrictive, and AI techniques were generally favored in the 1990s for their logical rather than mathematical approach [18,27,28,31]. The contribution of AI techniques is still significant for posing simulation-based design as a systematic problem, and for initiating the use of computational techniques to solve them. However, they have been shown to have very limited applicability.

Many improvements in numerical methods, solution strategies, and development of new algorithms over the past decade now allow a wide range of complex problems to be solved effectively. Building performance functions are often smooth in theory, and where possible, gradient-based methods have been shown to be very efficient and reliable [6,11,25,38]. However, use of complex simulations often results in derivative discontinuities, because of which gradient-based methods can fail.

Derivative-free deterministic methods such as generalized pattern search [1], DIRECT [12], and lattice methods [32] have been shown to perform particularly well for problems that suffer from simulation noise [26,29,34], but are generally limited to small problems. In addition to discontinuities in function-responses, simulation-based optimization can also be time-consuming since each design evaluation involves the use of simulation. Recent applications [34,6] have addressed both these issues by using approximation-based methods that derive simpler functions of the original simulation responses and use them for a partial search during the optimization process.

Over the past few years, stochastic methods such as simulated annealing and genetic algorithms have also become very popular, and have been applied to a range of problems for optimizing thermal and lighting performance based on building enclosure, HVAC design, and control schedules [3,5,7,37]. These methods are attractive mainly because they solve a wide range of problems, do not require functions to be smooth, and can handle mixed-discrete variables. However, these methods are based on random search and will often derive unreliable results unless used with considerable skill and intuition.

Hybrid strategies combining two or more methods have been used to overcome problems associated with one particular method. For example, Michalek et al. [19] and Monks et al. [22] use the global and versatile nature of stochastic methods with rigor and efficiency of gradient-based methods in a combined framework. [35] propose to combine genetic algorithms and pattern search [10] to derive a hybrid method that reduces computational run time in problems involving expensive simulations.

In addition to exploring a range of different methods, the applications from the past few years also demonstrate that with good understanding of the methods involved, design optimization can be effectively used to improve building performance and provide rigor in the way we use simulation tools. However, simulation-based design tasks often require expert analysis of different but interconnected performance goals. In principle it is desirable to evaluate interrelated design decisions concurrently so that their combined effects on different performance goals may be maintained. However, combining all design decisions and evaluating them simultaneously is difficult because it involves multiple and often conflicting performance goals, which may require expert analysis at very different levels of complexity and with different design information.

Past work on the simultaneous optimization of multiple performance goals has used multi-criteria formulations with preference or non-preference-based strategies. These applications typically provide the decision maker with values for decision variables that best accommodate a weighted set of performance criteria. The difficulty in elaborating them to include more than a few analytical tasks and decision variables is that the problem quickly becomes too large and complex to be implemented in one model. Even when numerical results are successfully obtained, one may not be able to interpret the design trade-offs or use intuition to confirm computed results [24]. For large and complex cases, some form of problem decomposition becomes necessary. In addition to making a problem manageable, decomposition of a large problem by focus or discipline is beneficial because it allows for the specialized analysis and decision making of individual design tasks. On the other hand, when the decision maker separates and designs individual parts of the problem, he must not only coordinate common decisions between different problems, but also combine the solutions into a single compatible set.

To coordinate design decisions conducted in separate parts, or for different performance aspects of the same problem, a rigorous and tractable process is necessary. This article thus presents a simulation-based design optimization model that: (a) is rigorous; (b) facilitates computational decision support for meeting performance goals; (c) allows autonomy of specialized design tasks with timely and efficient use of analysis tools; and, (d) preserves dependencies between various performance goals associated with a design, thereby enabling concurrent and consistent decision making.

2. Simulation-based building design by hierarchical optimization

Solving a combination of interrelated problems belongs to the realm of systems design, which is the branch of engineering concerned with the development of large and complex systems. Since simulation-based design problems require multiple criteria and diverse modes of analysis, they can be placed in the category of systems design, in which solving numerous performance goals for one problem is similar to assembling interconnected decision-making tasks. In this paper we present Analytic Target Cascading (ATC) in the context of thermal and HVAC design [4]. Hitherto used for automotive designs [13,14], ATC is a hierarchical optimization methodology for achieving compatible design targets in large engineering systems at early product development stages. It is based on the premise that the performance of a system element can be derived analytically as a function of its decision variables.

The following assumptions apply for using ATC as a methodology for handling thermal and HVAC design problems:

- Performance goals can be embodied as design targets that are to be achieved via design decisions. Some of the performance goals can be set as overall design targets, and introduced as part of initial problem definition.
- These targets, along with any required performance specifications, can be computed as functions of design decisions by using analysis/simulation models.
- A complex simulation-based design problem can be decomposed or partitioned into subproblems that can be further decomposed.
- It is possible to identify a hierarchical organization in the decomposition.

In the context of simulation-based design, a particularly beneficial feature of this decomposition–coordination approach is that each subproblem in the hierarchy constitutes a separate optimization problem and is associated with only those analysis models that are capable for computing the values of performance goals set for it. This allows both optimization algorithms and analysis tools to be used exclusively for the relevant decision-making problem.

The following section describes the ATC methodology. This methodology was earlier referred to by Michelena et al. [21] and more thoroughly presented by Kim [14]. While the ideas remain faithful to the original formulation, the general notation has since evolved. The notation followed in this article refers to [20] and [23]. Through a pilot application, we demonstrate how problem-specific selection of optimization methods can be used for solving simulation-based design tasks in buildings. Furthermore, application of optimization methods in building design is extended not only to computing efficient solutions for specific problems, but also to deriving compatible values of dependencies between multiple interconnected decision problems.

3. The ATC process

ATC is a multidisciplinary hierarchical optimization methodology that provides a systematic process for propagating desired top-down performance targets to appropriate lower level performance values (Fig. 1; [23]). In the ATC framework, the original design problem is partitioned into a set of subproblems constituting system, subsystems, and components. Design targets are specified at the top level of the multi-level design formulation and "cascaded down" to lower levels. Subproblems at lower levels are formulated so that all elements included in the hierarchy match the cascaded targets consistent with the overall system targets. Design targets derived at lower levels are rebalanced to higher levels by iteratively adjusting values of targets and decision variables.

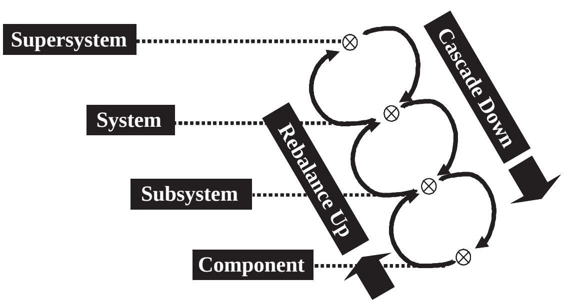

Fig. 1. Analytic Target Cascading.

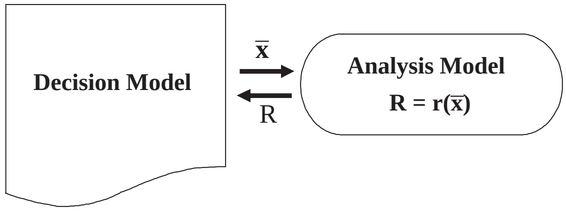

Fig. 2. Link between decision and analysis model in the ATC framework.

Each subproblem in the ATC hierarchy requires a decision model and one or more analysis models. The decision model of a subproblem is its formulation as a design optimization model. It requires representation of subproblem performance $\mathbf{R}$, decision variables $\bar{\mathbf{x}}$, and all relevant constraints $\mathbf{g}$ and $\mathbf{h}$ on decision variables. The decision model also embodies the links of each subproblem to upper and lower level subproblems in the hierarchy. It is through these links that top-level targets are propagated down and lower level responses are rebalanced up the hierarchy. Each decision model is associated with one or more analysis models to compute performance $\mathbf{R}$ as a function of decision variables $\bar{\mathbf{x}}$. The analysis model/s take(s) values of decision variables as input and returns their corresponding performance response as output (Fig. 2). Every analysis model requires an analysis tool (a simulation, for example) or an analytic function $\mathbf{r}$ from which performance R can be derived with respect to decision variables $\bar{\mathbf{x}}$. In the building simulation context, the simulation will typically return a data set (for example, a vector of room temperatures at every time step), which is processed by the analysis model into the required performance response (for example, maximum daily temperature).

Every subproblem in the ATC hierarchy is formulated and solved independently, and is posed to optimize "target matching" with its upper and lower level subproblems. The ATC problem is solved iteratively for meeting all targets as closely as possible by a coordination strategy. Once compatible targets are derived from the ATC process, individual subproblems can be isolated and outsourced to be solved in further detail, thereby enabling truly concurrent design [15]. In the ATC methodology, this constitutes the embodiment design step [14]. ATC is a generalizable design methodology with proven convergent properties [20]. Fig. 3 shows these basic steps in the ATC process, as well as the embodiment design step that follows it.

Formulating a design problem in the ATC framework requires: (a) identifying appropriate decomposition, (b) hierarchical organization of decomposed subproblems and identifying key links between them, (c) formulation of subproblems as decision models and identifying suitable optimization algorithms to solve them, and (d) building and mapping appropriate analysis models to each decision model.

Once formulated, steps involved in solving an ATC problem can be summarized as: (a) specifying values of overall design targets (referred to formally as "target setting"), (b) propagating specified top-level targets to lower levels and optimizing all subproblems to match targets as closely as possible, and (c) iteratively searching for an overall consistent solution by applying a coordination strategy.

3.1. Hierarchical formulation

The first step in setting up an ATC problem is to decompose the problem hierarchically. Typically four types of decomposition strategies are commonly found in systems design literature [33]: Object, Aspect, Sequential and Model decompositions. Object decomposition divides a system by physical components such as building, zones, rooms and components. Aspect decomposition divides the system according to different specialties or disciplines and is relevant when multiple performance aspects are evaluated for a physical component (for example, lighting and air distribution). Sequential decomposition is applicable when partitioned subproblems are organized by workflow or process logic. Sequential decomposition presumes uni-directionality of design information. Model decomposition is a more rigorous problem-partitioning method based on functional dependencies between decision variables and functions included in the problem.

By this method, unconnected or weakly connected structures in the overall decision model are identified by examining all included functions against decision variables [24]. Such decompositions can be applied to simulation-based design problems by notions of physical hierarchy in buildings, by the type of design information an analysis or simulation requires, by performance aspects included in the model, or by logic of information flow.

Once the decision model is decomposed, appropriate analysis models are associated with each

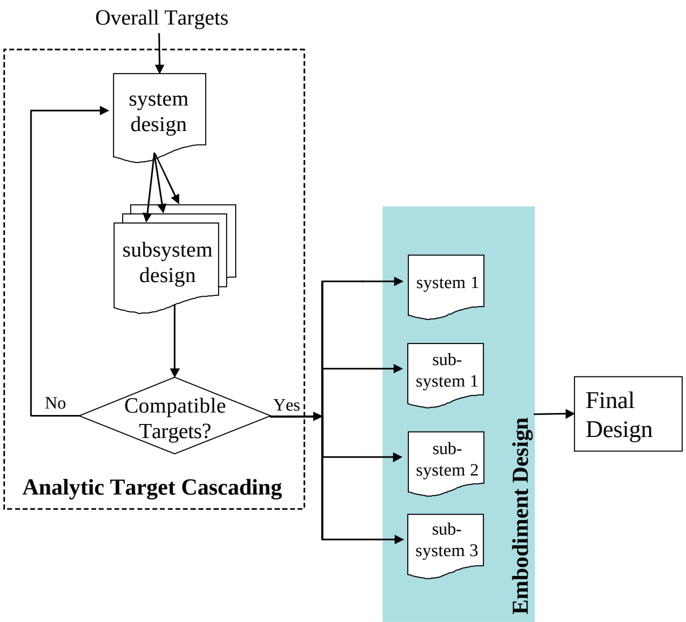

Fig. 3. Basic steps in the ATC process.

element in the hierarchy. An analysis model is appropriate if it can compute performance values as functions of its decision variables. The existence of appropriate analysis models is presumed in this setup. Every analysis model evaluates design decisions by taking variables and parameters as input and returning performance values as output.

In the ATC framework, problem partitioning also includes identifying common links between subproblems. Horizontal links between subproblems are called linking variables. Linking variables represent decisions shared among two or more decision models at the same vertical level [14]. Their value is determined individually by the decision models that share them and coordinated by the upper level parent problem. This means that subproblems that share linking variables must also have a common decision model at the upper level.

The vertical relationships between decomposed levels are embodied by performance targets and responses. Fig. 4 shows the information flow up and down the ATC hierarchy for a three-level problem with one system, three subsystems, and two component level problems. Rectangular boxes represent the decision models, and oval boxes are the analysis models. As an example, the subsystem B model receives its target values $\mathbf{R}_\mathrm{B}^\mathrm{U}$ from the system level model S. At an iteration k in the ATC process, subsystem B also has target values of responses $\mathbf{R}_\mathrm{B1}^\mathrm{L}$ and $\mathbf{R}_\mathrm{B2}^\mathrm{L}$, and linking variables $\mathbf{y}_\mathrm{B1}^\mathrm{L}$, and $\mathbf{y}_\mathrm{B2}^\mathrm{L}$ from the component levels. Subsystem B is solved for determining values of its local decision variables $\tilde{\mathbf{x}}_\mathrm{B}$, values of component level responses $\mathbf{R}_\mathrm{B1}$ and $\mathbf{R}_\mathrm{B2}$, and values of coordinating linking variables $\mathbf{y}_\mathrm{B}$ such that deviations from information received from upper and lower levels are minimized. This includes minimizing deviations between $\mathbf{R}_\mathrm{B1}$ and $\mathbf{R}_\mathrm{B1}^\mathrm{L}$, $\mathbf{R}_\mathrm{B2}$ and $\mathbf{R}_\mathrm{B2}^\mathrm{L}$, $\mathbf{y}_\mathrm{B}$ and $\mathbf{y}_\mathrm{B1}^\mathrm{L}$, $\mathbf{y}_\mathrm{B}$ and $\mathbf{y}_\mathrm{B2}^\mathrm{L}$, and $\mathbf{R}_\mathrm{B}$ and $\mathbf{R}_\mathrm{B}^\mathrm{U}$. Subsystem B computes its response $\mathbf{R}_\mathrm{B}$ by using its analysis model. It provides values of local decision variables $\tilde{\mathbf{x}}_\mathrm{B}$ and lower level responses $\mathbf{R}_\mathrm{B1}$ and $\mathbf{R}_\mathrm{B2}$ as input to the analysis model. The analysis model returns the subsystem response $\mathbf{R}_\mathrm{B}$ as output. Note that in these interactions responses of a particular level are decision

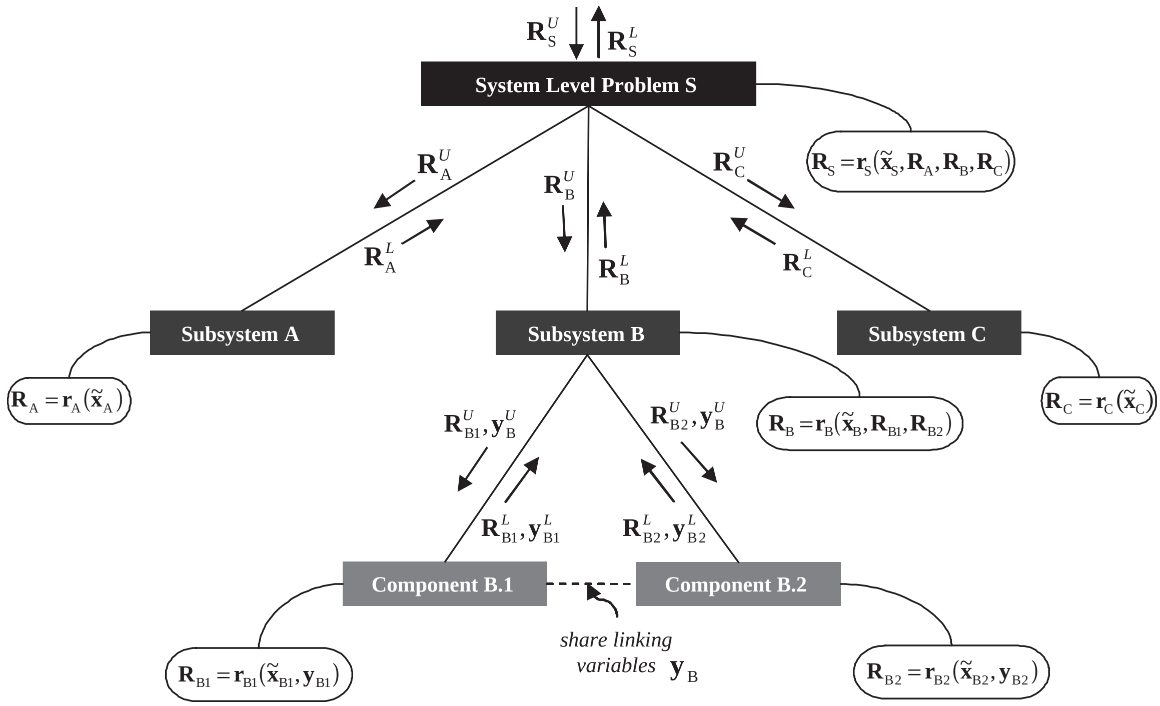

Fig. 4. Information flow up and down the ATC hierarchy (figure shows a three-level example problem with one system, three subsystems, and two component level problems).

variables input to the analysis model at the level above. Another condition in this organization is that linking variables are shared between children of a parent problem.

3.2. Mathematical models

Fig. 5 shows the ATC formulation in the standard index notation given by Michelena and Park [20]. For target matching a problem S_{ij} for the j-th design model at the i-th level, the general formalization of the optimization model is stated as:

$$\underset{\bar{\mathbf{x}}_{ij},\mathbf{y}_{(i+1)j},\varepsilon^R_{ij},\varepsilon^y_{ij}}{\text{minimize}} \left\|\mathbf{R}_{ij}-\mathbf{R}^{\mathrm{U}}_{ij}\right\|^2_2+\left\|\mathbf{y}_{ij}-\mathbf{y}^{\mathrm{U}}_{ij}\right\|^2_2+\varepsilon^R_{ij}+\varepsilon^y_{ij} \tag{1}$$

subject to:

$$\sum_{k\in C_{ij}}\left\|\mathbf{R}_{(i+1)k}-\mathbf{R}^{\mathrm{L}}_{(i+1)k}\right\|^2_2\le\varepsilon^R_{ij},$$

$$\sum_{k\in C_{ij}}\left\|\mathbf{y}_{(i+1)k}-\mathbf{y}^{\mathrm{L}}_{(i+1)k}\right\|^2_2\le\varepsilon^y_{ij},$$

$$\mathbf{g}_{ij}\left(\bar{\mathbf{x}}_{ij}\right)\le\mathbf{0},$$

$$\mathbf{h}_{ij}\left(\bar{\mathbf{x}}_{ij}\right)=\mathbf{0},$$

where:

- $\bar{\mathbf{x}}_{ij}$=[$\tilde{\mathbf{x}}_{ij}$, $\mathbf{y}_{ij}$, $\mathbf{R}_{(i+1)k_1}$,. . .,$\mathbf{R}_{(i+1)k_{c_{ij}}}$]^T is the vector of all decision variables of element j at level i,
- $\mathbf{R}_{ij}$=$\mathbf{r}_{ij}(\bar{\mathbf{x}}_{ij})$ where $\mathbf{r}_{ij}$ is the vector function that represents the analysis model. It calculates the responses for element j at level i by taking in all its decision variables as input.
- C_{ij}={k_1,. . .,$k_{c_{ij}}$}, and c_{ij} is the number of child elements,

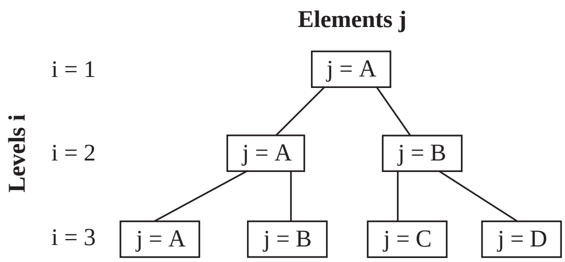

Fig. 5. Standard index notation for a hierarchically partitioned problem.

- $\tilde{\mathbf{x}}_{ij}\in\Re^{n_{ij}}$ is the vector of local decision variables for element j at level i,
- $\mathbf{y}_{ij}\in\Re^{l_{ij}}$ is the vector of linking variables for element j at level i,
- $\mathbf{R}_{ij}\in\Re^{d_{ij}}$ is the vector of local responses for element j at level i. It is a function of local, linking, and children response variables: $\mathbf{r}_{ij}:\Re^{n_{ij}+l_{ij}+\sum_k d_{(i+1)k}}\rightarrow\Re^{d_{ij}}$,
- ε^R_{ij} is the tolerance variable for consistency of targets set at element j level i and the responses of j's children,
- ε^y_{ij} is the tolerance variable for consistency of linking variables coordinated at element j level i for child elements at the (i+1)*th* level,
- $\mathbf{R}^{\mathrm{U}}_{ij}\in\Re^{d_{ij}}$ is the vector of response values cascaded to element j at level i as targets from its parent at level ($i-1$),
- $\mathbf{y}^{\mathrm{U}}_{ij}\in\Re^{l_{ij}}$ is the vector of coordinating linking variables for the linking variables in the children of element j at level i. This vector includes one copy of each linking variable from all of element j's children.
- $\mathbf{R}^{\mathrm{L}}_{(i+1)k}\in\Re^{d_{(i+1)k}}$ is the vector of response variable values cascaded to the element j at level i from its k-th child at level ($i-1$),
- $\mathbf{y}^{\mathrm{L}}_{(i+1)k}\in\Re^{l_{(i+1)k}}$ is the vector of linking variable values cascaded to the element j at level i from its $k-$th child at level ($i-1$),
- $\mathbf{g}_{ij}:\Re^{d_{ij}+n_{ij}+l_{ij}}\rightarrow\Re^{v_{ij}}$ and $\mathbf{h}_{ij}:\Re^{d_{ij}+n_{ij}+l_{ij}}\rightarrow\Re^{s_{ij}}$ are vector functions representing inequality and equality design constraints,
- $\|\cdot\|^2_2$ represents the square of the l_2 norm.

In this general formulation, the overall system targets are specified (represented as $\mathbf{R}_{ij}$=$\mathbf{T}$). Also, subproblems that do not have any children will not contain the tolerance variable for coordinating lower level information. Subproblems that do not share any linking variables with other children of their parent element will not have the terms for minimizing linking variable deviations, and parents whose children do not share any linking variables will not include the tolerance variable for coordinating linking variables.

3.3. Hierarchical coordination

The order of solving each decision model in the ATC hierarchy and dispatching its solution to other models in the hierarchy is called coordination strategy. Once the ATC problem is set up for all elements in a hierarchy in the form shown in Eq. (1), a coordination strategy is applied to iterate through the multi-level structure. The main requirement for a coordination strategy is that it should converge to the same solution set as that of an unpartitioned problem [20,24]. A general assumption in the ATC literature is that the optimization problems are continuous. However the formulation is also valid for mixed-discrete problems.

The ATC formulation is both top-down and bottom-up, which means that once some overall performance targets are specified at the topmost level, they are disseminated down to determine the value of lower level performances. Likewise, if performance targets are initially specified for lower level problems, then the upper level performances can be determined based on the information propagated up from the lower levels. In this manner, the solutions derived at any particular level in the hierarchy are sought to be consistent with all other components of the partitioned problem. The iterative process of cascading targets is repeated until specified termination criteria are met.

3.4. Relevance and benefits to consistent simulation-based design

In addition to the general benefits of problem-partitioning and explicit means of deriving system optimality, ATC offers a goal-driven method in contrast to the data-centric approaches to collaborating design decisions in building simulation. It is a process in which dialogues between multiple decision making tasks are singularly motivated towards meeting performance specifications. Instead of direct transfer/sharing of information among various simulations from a data repository, this framework connects simulations via decision-making models, or in other words, through the purposes for which they are required (see Fig. 4). Communications among simulations thus occur only where necessary. This is a particularly beneficial prospect in the building design context where communications between interrelated simulations can be overwhelming and sometimes even indiscernible. It is also an approach that is likely to be more rewarding because all aspects involved in a simulation-based scenario including making decisions, evaluating trade-offs, or invoking a particular analysis tool, depend on what performance specifications of a particular problem. Likewise, their consideration is central to the ATC model.

In addition, the ATC process seeks solutions that are compatible with all performance goals included in the problem. If a compatible set is not feasible, the ATC process allows the modeler to examine trade-offs explicitly. Explicit knowledge of trade-offs between performance targets can better inform decision making in practical settings where performance goals are likely to be contradictory.

The clear separation between the decision and the analysis model is also an important feature in this hierarchical decomposition approach. It implies that the decision models can be freely modified as per the problem, by adding or changing performance goals or decision variables. New functional relations can be included as additional constraints depending on the peculiarities of the problem under consideration. Furthermore, analysis tools can be used efficiently without information overload of irrelevant decision variables. The independence of the decision model from the analysis model also provides the flexibility to use any relevant simulation tool that is available. Finally, for any particular subproblem in the hierarchy, multiple analysis models can be also used if required. The following section extends this methodology to the building design context through a pilot application.

4. Pilot application

Thermal and HVAC design illustrate a typical simulation-based problem outsourced to specialists by specifying a set of performance goals and design decisions by which those goals much be achieved. Such problems permit partitioning decision-making tasks into subproblems defined by the design decisions involved, the type of design information a particular analysis or a simulation requires, the performance aspect, or the physical zone being considered.

The following problem extends the ATC framework to the context of simulation-based design in buildings by applying it to the thermal design and analysis of a fictional design scenario. This design case also serves to validate the multi-level optimization approach against solving the problem altogether in one model by comparing results obtained from both. Symbols used throughout this study are given at the end of this article (Table 8).

4.1. Design scenario

Thermal design of a three-zone building consisting of one office and two workshops is considered (Fig. 6). This problem is defined for making decisions on the sizes of walls and windows, thermostatic set point temperatures, and average zone velocities for meeting specified performance targets. The two workshops (zones 2 and 3) are assumed to be identical in terms of use, environmental conditions, and occupancy.

Depending on the function and schedules of use, the following targets are specified for the building:

(a) Overall floor area: while seeking the optimum wall sizes of each zone, the total building area A, must match the given overall area T_A of the building.
(b) Total thermal performance of zone 1, 2 and 3: 100% "thermal performance" is targeted for the building. This implies that at the best case, the summation of the thermal performance P of all three zones should match target T_P=1.

Area A and thermal performance P represent the overall building responses, and are derived by summing individual zone area and performance:

$$A = 2A_w + A_o, \text{ and } P = (2P_w + P_o)/3 \tag{2}$$

where A_w and A_o are the areas of the workshop and office respectively, and P_w and P_o are their thermal performance values. Thermal performance P_w of the workshop is formulated for minimizing annual heating and cooling energy. So P_w=1 when the total energy consumption of the workshop is zero. Thermal performance P_o of the office is defined for maximizing the number of occupied hours over which a given occupant comfort index is maintained, and P_o=1 when the average absolute value of thermal comfort index PMV over all occupied hours is zero. Since the two workshops are identical, their area and thermal performance is computed for one zone and doubled to include the other.

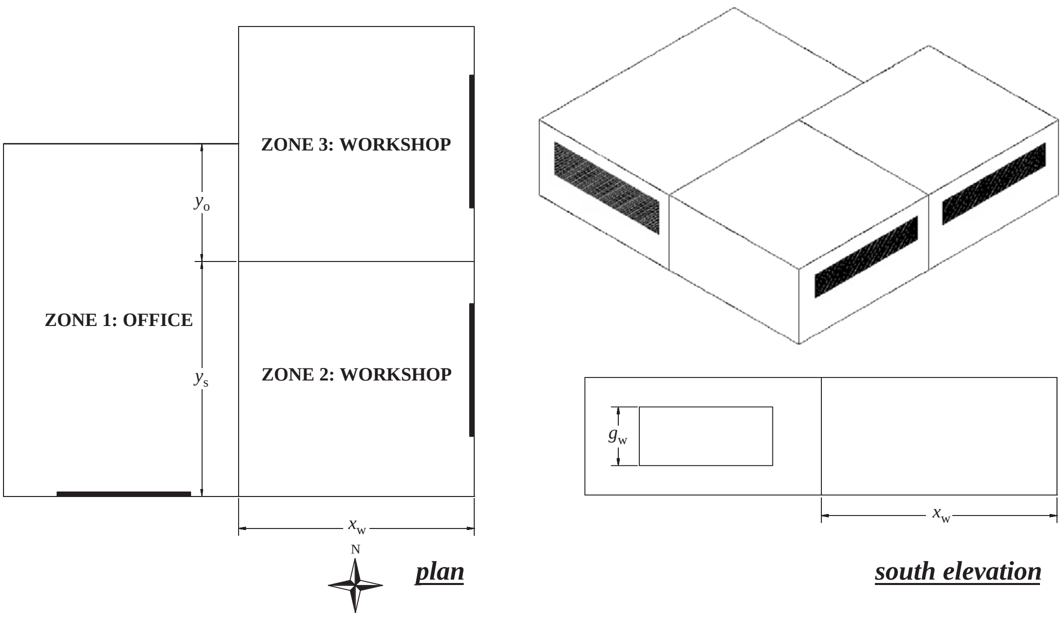

Fig. 6. The Three Zone Example Case.

The decision variables included in this problem are: (a) width of the workshop x_w, (b) y_s and y_s+y_o as length of the workshop and office respectively, (c) window wall percentage of the workshop g_w, (d) window height of the office g_o, (e) mean zone velocity v_w and v_o, heating set point temperatures t_w^h and t_o^h, and cooling set point temperatures t_w^c and t_o^c of the workshop and office, respectively.

A_w and P_w are derived from the decision variables and parameters related to the workshops, and A_o and P_o are functions of the office attributes only except for y_s, which is the length of the common wall between the office and the workshop.

In addition to meeting overall targets T_A and T_P, each zone design must be within defined feasible limits of average room temperatures and thermal comfort at all occupied hours and total annual energy cost.

4.2. ATC formulation: hierarchical decomposition

The problem as posed can be decomposed in a bilevel structure, based on physical distinctions between the zones (see Fig. 7). The overall thermal design problem is represented at the system level where target values T_A and T_P are specified for building area and thermal performance. The problem is decomposed into two subproblems at the second level. Subsystem 1 represents the office and subsystem 2 represents the workshops.

The system level problem **Z** determines optimal values of zone area and thermal performance (A_w, A_o and P_w, P_o) for meeting specified building targets T_A and T_P. Values of zone area and performances derived at this level are cascaded down to the appropriate zone level problem as targets. The subsystem problems are solved for meeting their area and thermal performance targets that are passed down from the upper level, while satisfying design constraints. The common wall between the workshop and office (y_s) is represented as a linking variable, which means that it is coordinated by the system level problem **Z**. The following mathematical formulations are used to set up the ATC process for this problem. Symbols used throughout this example are also given in Table 8.

4.2.1. System level model

The system level problem is posed for meeting specified targets T_A=597 m^2 and T_P=1 with respect to zone areas A_w, A_o and thermal performance P_w, P_o, for minimizing deviations between values of zone

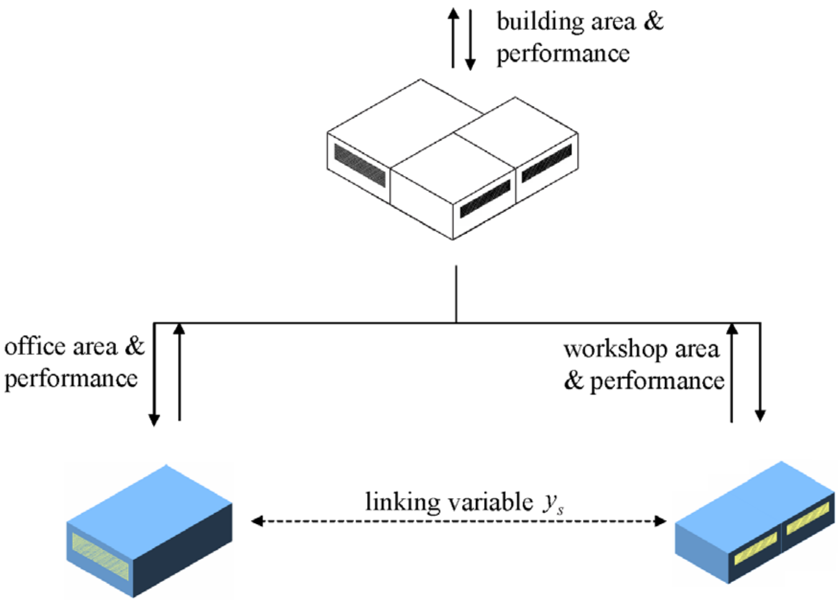

Fig. 7. Decomposition of the Three-Zone Thermal Design Problem.

area and performance derived at this level with those passed back from the subsystem level, and for coordinating the value of linking variable shared between the two subsystem problems. The optimization model for the system level problem is formally stated as:

$$\mathbf{Z}: \text{ Minimize } (P - T_{\mathrm{P}})^2 + (P - T_{\mathrm{A}})^2 + \varepsilon_{\mathrm{R}} + \varepsilon_{\mathrm{y}} \tag{3}$$

with respect to:

$$\mathbf{R}_{\mathrm{Zs}} = (A_{\mathrm{w}}, A_{\mathrm{o}}, P_{\mathrm{w}}, P_{\mathrm{o}})$$

$$\mathbf{y}_{\mathrm{Zs}} = (y_{\mathrm{s}})$$

$$\varepsilon_{\mathrm{R}}, \varepsilon_{\mathrm{y}} = (\varepsilon_{\mathrm{R1}}, \varepsilon_{\mathrm{R2}}, \varepsilon_{\mathrm{R3}}, \varepsilon_{\mathrm{R4}}, \varepsilon_{\mathrm{y}})$$

where:

$$\mathbf{R}_{\mathrm{Z}}[P, A] = r_{\mathrm{Z}}(\mathbf{R}_{\mathrm{Zs}})$$

subject to:

$$(A_{\mathrm{w}} - A_{\mathrm{w}}^{\mathrm{L}})^2 \le \varepsilon_{\mathrm{R1}} \quad (A_{\mathrm{o}} - A_{\mathrm{o}}^{\mathrm{L}})^2 \le \varepsilon_{\mathrm{R2}}$$

$$(P_{\mathrm{w}} - P_{\mathrm{w}}^{\mathrm{L}})^2 \le \varepsilon_{\mathrm{R3}} \quad (P_{\mathrm{o}} - P_{\mathrm{o}}^{\mathrm{L}})^2 \le \varepsilon_{\mathrm{R4}}$$

$$\frac{1}{2}\left(\left(y_{\mathrm{s}} - (y_{\mathrm{s}}^{\mathrm{L}})_{\mathrm{Zs1}}\right)^2 + \left(y_{\mathrm{s}} - (y_{\mathrm{s}}^{\mathrm{L}})_{\mathrm{Zs2}}\right)^2\right) \le \varepsilon_{\mathrm{y}}$$

$$A_{\mathrm{w}}^{\min} \le A_{\mathrm{w}} \le A_{\mathrm{w}}^{\max} \quad A_{\mathrm{o}}^{\min} \le A_{\mathrm{o}} \le A_{\mathrm{o}}^{\max}$$
$$P_{\mathrm{w}}^{\min} \le P_{\mathrm{w}} \le P_{\mathrm{w}}^{\max} \quad P_{\mathrm{o}}^{\min} \le P_{\mathrm{o}} \le P_{\mathrm{o}}^{\max}$$
$$y_{\mathrm{s}}^{\min} \le y_{\mathrm{s}} \le y_{\mathrm{s}}^{\max}$$

In this particular example there are no local variables at the system level (see Table 1). Therefore, the only constraints are those formulated for coordinating the information received from the lower level. In addition to minimizing difference between targets and performances, the deviation tolerances (ε_{R} and ε_{y}) used for coordinating lower level information are also minimized. In an ideal case the deviation tolerances are expected to be zero, implying that values of the building level variables match the performance values determined at the level below, and the value of the shared variable y_s is set the same in both subsystems. All objectives in this multi-criteria formulation are equally weighted, and all targets and responses are scaled between 0 and 1.

Table 1
Summary of responses and variables at the system level

System level problem	**Z**
System responses ($\mathbf{R}_{\mathrm{Zs}}$)	$[P, A]$
Local variables ($\tilde{\mathbf{x}}$)	–
Responses from the subsystem level ($\mathbf{R}_{\mathrm{Zs}}$)	$[P_{\mathrm{o}}, P_{\mathrm{w}}, A_{\mathrm{o}}, A_{\mathrm{w}}]$
Subsystem level linking variables ($\mathbf{y}_{\mathrm{Zs}}$)	$[y_{\mathrm{s}}]$

4.2.2. Subsystem level models

The optimization problems at the subsystem level are posed for finding values of local variables and linking variable such that the deviations between target zone area and performance cascaded down from the upper level are minimized, while satisfying all local constraints. Since both subsystem problems 1 and 2 represent the last levels in this example, their problem formulation will not include any terms for coordinating of lower level information.

The local decision variables included in subsystem 1 (office decision model) are: (a) partial length of the office y_{o}, (b) window height of the office, (c) mean zone velocity v_{o}, (d) heating set point temperature $t_{\mathrm{o}}^{\mathrm{h}}$, and (e) cooling set point temperature $t_{\mathrm{o}}^{\mathrm{c}}$. Subsystem 2 (workshop decision model) local variables include: (a) width of the workshop x_{w}, (b) window wall percentage of the workshop g_{w}, (c) mean zone velocity v_{w}, (d) heating set point temperature $t_{\mathrm{w}}^{\mathrm{h}}$, and (e) cooling set point temperature $t_{\mathrm{w}}^{\mathrm{c}}$. Length of the common wall between the office and workshop y_{s} is the linking variable.

The scaled area of the office A_{w} and the workshop A_{o} are derived from the following decision variables and parameters (see Fig. 6):

$$A_{\mathrm{w}} = k_{\mathrm{w}}(x_{\mathrm{w}} y_{\mathrm{s}}) \tag{4}$$

$$A_{\mathrm{o}} = k_{\mathrm{o}}(x_{\mathrm{o}}(y_{\mathrm{s}} + y_{\mathrm{o}}))$$

where x_{w} and x_{o} are the width of the workshop and the office respectively, y_{s} is the length of the workshop and also the shared wall between the office and the workshop, and y_{s}+y_{o} is the length of the office. k_{w} and k_{o} are the scaling parameters for A_{w} and A_{o}, respectively.

Thermal performance P_{w} and P_{o} are evaluated with respect to the decision variables by using a simulation tool that takes the values of the decision variables as input, and returns their corresponding performance values as responses. P_w represents the sum of total annual heating and cooling energy load of the office scaled by the maximum feasible amount of energy load E_{refw}:

$$P_w = \left(\sum_{i=1}^{m} H_i^w + \sum_{i=1}^{m} C_i^w \right) \Big/ E_{refw} \quad (5)$$

where H_i^w and C_i^w are the heating and cooling energy at time step i (i=60 min), and m is the total number of time steps in the year for which H_i^w and C_i^w are computed. Thermal performance P_o is a thermal comfort index representing the average absolute value of thermal comfort in the office (measured in PMV) over all occupied hours $\bar{m}$:

$$P_o = \left(\sum_{i=1}^{\bar{m}} \left| \varphi_i^o \right| \right) \Big/ \bar{m} \quad (6)$$

In addition to meeting overall targets T_A and T_P, each zone design must be within defined feasible limits of average room temperatures, thermal comfort, and total annual energy cost. These thermal performance indices are also computed with respect to the decision variables included in the problem.

Both office and workshop are required to maintain average air temperature between 20 °C and 23 °C during all occupied hours $\bar{m}$. The analysis model evaluates the zone average air temperatures at all $\bar{m}$ and returns the maximum and minimum air temperature of each zone to the decision model, which evaluates them by the following statements:

$$\begin{array}{ll} t''_w \le t''_{ref}, & t''_o \le t''_{ref} \\ t'_w \le t'_{ref}, & t'_o \le t'_{ref} \end{array} \quad (7)$$

where t''_w and t''_o are the maximum mean air temperatures of the workshop and office respectively, t'_w and t'_o are the minimum mean air temperatures of the workshop and office respectively, t''_{ref} is the maximum allowed temperature of the zones (23 °C), and t'_{ref} is the minimum allowed temperature of the zones (20 °C).

Absolute value of the thermal comfort index (measured in PMV) at every occupied hour is required to be less than a specified upper limit. The analysis model returns the maximum absolute value of thermal comfort over all $\bar{m}$ to the decision model, which evaluates it against the maximum feasible value:

$$\varphi''_w \le \varphi_{ref}, \text{ and } \varphi''_o \le \varphi_{ref} \quad (8)$$

where φ''_w and φ''_o are the maximum thermal comfort values of the workshop and office respectively, and φ_{ref} is their maximum feasible value.

The total annual heating and cooling energy of both zones (E_w and E_o) is also constrained by their specified upper limits:

$$E_w \le E_{refw}, \text{ and } E_o \le E_{refw} \quad (9)$$

where $E_w = \left(\sum_{i=1}^{m} H_i^w + \sum_{i=1}^{m} C_i^w \right)$, and $E_o = \left(\sum_{i=1}^{m} H_i^o + \sum_{i=1}^{m} C_i^o \right)$. E_{refw} and E_{refo} are specified upper limits for E_w and E_o, respectively.

The mathematical decision models of the zone and workshop are stated as follows:

Subsystem 1: The Office Decision Model

$$\mathbf{Z}_{s1}: \ \text{Minimize}\left(P_o - P_o^U\right)^2 + \left(A_o - A_o^U\right)^2 + \left(y_s - y_s^U\right)^2 \quad (10)$$

with respect to:

$$\tilde{\mathbf{x}}_o = \left(y_o, g_o, v_o, t_o^h, t_o^c\right)$$

$$\mathbf{y}_{Zs1} = (y_s)$$

where:

$$\mathbf{R}_{Zs1}[P_o, A_o, E_o, t''_o, t'_o, \varphi''_o] = \mathbf{r}_{Zs1}(\tilde{\mathbf{x}}_o, \mathbf{y}_{Zs1})$$

subject to:

$$\begin{array}{ll} t''_o - t''_{ref} \le 0 & t'_{ref} - t'_o \le 0 \\ \varphi''_o \le \varphi_{ref} & E_o \le E_{refo} \end{array}$$

$$\begin{array}{ll} y_o^{min} \le y_o \le y_o^{max} & g_o^{min} \le g_o \le g_o^{max} \\ v_o^{min} \le v_o \le v_o^{max} & \\ \left(t_o^h\right)^{min} \le t_o^h \le \left(t_o^h\right)^{max} & \left(t_o^c\right)^{min} \le t_o^c \le \left(t_o^h\right)^{max} \end{array}$$

Subsystem 2: The Workshop Decision Model

$$\mathbf{Z}_{s2}: \ \text{Minimize}\left(P_w - P_w^U\right)^2 + \left(A_w - A_w^U\right)^2 + \left(y_s - y_s^U\right)^2 \quad (11)$$

with respect to:

$$\tilde{\mathbf{x}}_w = \left(y_w, g_w, v_w, t_w^h, t_w^c\right)$$

$$\mathbf{y}_{Zs2} = (y_s)$$

where:

$$\mathbf{R}_{Zs2}[P_w, A_w, E_w, t''_w, t'_w, \varphi''_w] = \mathbf{r}_{Zs2}(\tilde{\mathbf{x}}_w, \mathbf{y}_{Zs2})$$

subject to:

$$t''_w - t''_{ref} \le 0 \quad t'_{ref} - t'_w \le 0$$
$$\varphi''_w \le \varphi_{ref} \quad E_w \le E_{refw}$$

$$y_w^{min} \le y_w \le y_w^{max} \quad g_w^{min} \le g_w \le g_w^{max}$$
$$v_w^{min} \le v_w \le v_w^{max}$$
$$\left(t_w^h\right)^{min} \le t_w^h \le \left(t_w^h\right)^{max} \quad \left(t_w^c\right)^{min} \le t_w^c \le \left(t_w^c\right)^{max}$$

4.3. Implementation setup

Sequential Quadratic Programming (SQP) is used to optimize of the system level performance targets. SQP is a gradient-based optimization algorithm, which means that it uses function gradients to make decisions about which designs to explore. The algorithm is fast for small problems and produces locally optimal design. Since SQP requires all decision variables and design relations to be smooth, superEGO [30] was used to optimize zone level problems where some design relations are non-smooth because of simulation "noise." An approximation-based global optimization algorithm, superEGO is also efficient for problems where the analysis models may include calls to expensive simulation tools. The algorithm takes an initial data sample of the objective function and fits a surrogate model to that data. It uses the surrogate model to search for optimal solutions and therefore reduces the number of calls to the analysis model. Although not highly critical in this demonstration case, this feature is an important consideration for any future applications where expensive simulations could make the ATC process extremely time-consuming.

All decision variables, constraints, and functions used in this formulation are scaled between 0 and 1. Table 2 summarizes the responses and variables included in the subproblem level models.

The analysis models of both subproblems use Energy Plus, a building energy analysis tool [8], for computing zone responses for evaluating targets as well as design constraints. All evaluations are run hourly, for 1 year. The analysis model aggregates the

Table 2
Summary of responses and variables at the subsystem level

Subsystem level problems	Zs1	Zs2
Subsystem responses ($\mathbf{R}_{Zs}$)	$[P_o, A_o]$	$[P_w, A_w]$
Local variables ($\tilde{\mathbf{x}}$)	$[y_o, g_o, v_o, t_o^h, t_o^c]$	$[y_w, g_w, v_w, t_w^h, t_w^c]$
Subsystem linking variables ($\mathbf{y}_{Zs}$)	$[y_s]$	$[y_s]$

thermal performance indices calculated by the simulation tool in appropriate forms and returns them to the decision model as design responses.

4.4. Hierarchical optimization

Fig. 8 shows the information flow up and down the levels, as well as input and output between the decision and analysis models. Starting from the system level, deviations between specified targets $[T_A, T_P]$ and system level responses $[A, P]$ are minimized with respect to subsystem responses $\mathbf{R}_{Zs1}$ and $\mathbf{R}_{Zs2}$. Values of $\mathbf{R}_{Zs1}$ and $\mathbf{R}_{Zs2}$ computed at the system level are cascaded down to corresponding subsystem problem as zone level targets $\mathbf{R}_{Zs1}^U$ and $\mathbf{R}_{Zs1}^U$. At this point, an estimated initial value of the linking variable $\mathbf{y}_{Zs}$ is also passed down to both zone level problems as target $\mathbf{y}_{Zs}^U$. Both subsystem level problems are now solved for values of local and linking variables such that: (a) deviations between zone level targets and linking variables cascaded from the upper level ($\mathbf{R}_{Zs1}^U$, $\mathbf{R}_{Zs1}^U$, and $\mathbf{y}_{Zs}^U$) are minimized, (b) zone responses computed at the subsystem level are minimized, and (c) all constraints are satisfied.

When the subsystem level problems are solved, the responses and linking variables computed by the zone level models are passed back to the system level model as lower level responses ($\mathbf{R}_{Zs}^L$, $\mathbf{R}_{Zs}^L$, and $\mathbf{y}_{Zs}^L$), and the building level problem is resolved with respect to zone responses and linking variables for matching overall targets and for minimizing deviations between zone area, performance, and linking variables passed back from the lower level.

When responses and linking variables computed at the subsystem level are passed back to the system level decision model, one target cascading iteration is complete. For this top-down case, the iterations were terminated when the deviation terms ε_R and ε_y

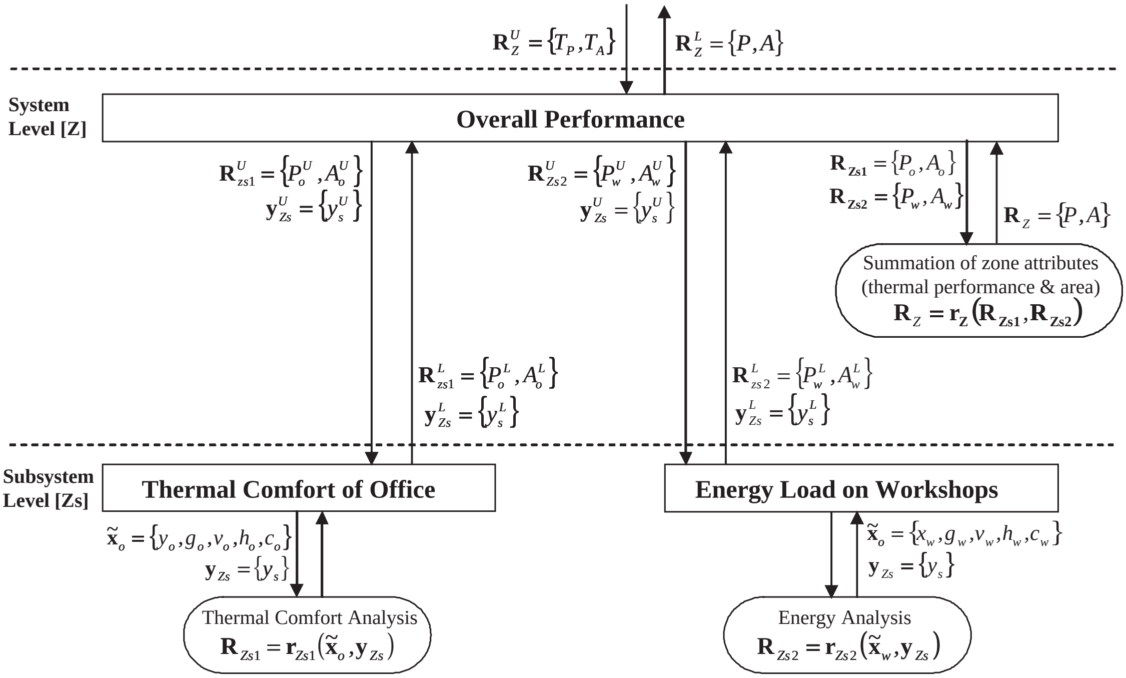

Fig. 8. Information Exchange between the Bi-level Thermal Design Case.

became smaller than a specified tolerance and also when the decision variables of all the models in the hierarchy stopped changing between subsequent iterations. At the end of the ATC process, all performance targets (T_A, T_P, A_o, A_w, P_o, P_w) and linking variables (y_s) are determined in compatibility with one another, i.e., meeting the targets as closely as possible, while satisfying constraints throughout the hierarchy.

4.5. ATC results

For the three-zone case, the ATC process terminated in 12 iterations. Table 3 shows the trade-offs between overall building and zone level targets. For the goal of maximizing thermal performance and area, these trade-offs represent the best compromise between performance targets while meeting all feasibility constraints. Table 4 shows the values of subproblem responses and coordinating linking variable determined at the top level. These values are passed down as targets to the subsystem levels, and Table 5 shows the subsystem solution for meeting these targets, and Table 6 shows the optimal values for subsystem level local variables. As shown in Table 5, the workshop targets do not match their target values. This is because the thermal performance target for the

Table 3
Overall system level targets and responses

Top-level model: overall building performance	Targets	Responses
Thermal performance of building P (%)	100	92.6
Overall area of building A (m^2)	597	521
Deviation between subsystem targets and responses ε_R (scaled)	0	0.05
Deviation among subsystem linking variables ε_y (scaled)	0	0.0025

Table 4
Top-level design solution

Top-level decision variables	Initial values	Optimal values	Lower bounds	Upper bounds
Thermal performance of office P_o (%)	1	1	0	1
Area of office A_o (m^2)	237	225	112	237
Thermal performance of workshop P_w (%)	1	0.89	0	1
Area of workshop A_w (m^2)	180	168	48	180
Linking variable y_s (m)	15	14	8	15

Table 5
Subsystem level targets and responses

Subsystem design solution	Targets	Responses
Subsystem 1: office model		
Thermal performance of office P_o (%)	1	1
Area of office A_o (m^2)	225	225.36
Linking variable y_s (m)	14	14.47
Subsystem 2: workshop model		
Thermal performance of workshop P_w (%)	0.89	0.94
Area of workshop A_w (m^2)	168	138
Linking variable y_s (m)	14	13.7

workshop (minimizing total energy) is unattainable, and while minimizing the energy cost, the area of the workshop is also reduced.

4.6. Validation of ATC results against "all at once" solution

Since this example case is small, these results are also compared to results obtained when the same problem was posed as one optimization problem, and solved "all at once," namely without the ATC decomposition. (This comparison is a validation of ATC given by Kim [14]). The "all at once" problem was optimized using superEGO. Table 7 shows both the ATC and "all at once" solution. The objective function value is the total deviation between all targets and performance goals, and A_o, P_o, A_w, and P_w are the performance values determined for the two zones.

Ideally, the solution to the "all at once" problem should be the same or better than the ATC results for the same problem. However, better results have been derived by the ATC process. The use of SQP at the system level of the ATC formulation allowed better convergence towards the true optimum. Additionally, partitioning of the problem results in lower dimensionality of optimization problem. It includes 5 variables per subproblem as against 11 variables when the problem is solved "all at once," and this increases the performance of superEGO and also improves convergence.

In general, computational expense of the analysis and non-smooth functional dependencies are common features of the energy analysis problem. Therefore, it is difficult to solve the whole problem using gradient-based optimization algorithms such as SQP. On the other hand, solving the whole problem using superEGO compromises the solution. As it turns out, by allowing the use of different optimization strategies suitable at specific levels of the system decomposition, the ATC process results in a better solution.

Table 6
Subsystem level decision variables

Subsystem variables	Initial values	Optimal values	Lower bounds	Upper bounds
Subsystem 1 (office model): Variables				
Partial length of office y_o (m)	4.75	4.31	1.3	4.75
Window height g_o (m)	1.0	0.84	0.5	1.5
Mean zone velocity v_o (m/s)	0.35	0.14	0.1	0.5
Heating set point temperature t_o^h (°C)	22	21.21	19	24
Cooling set point temperature t_o^c (°C)	22	21.34	19	24
Subsystem 1 (workshop model): Variables				
Width of workshop x_w (m)	12	10.1	6	12
Window wall percentage g_w (%)	0.8	0.67	0.2	0.8
Mean zone velocity v_w (m/s)	0.35	0.18	0.1	0.5
Heating set point temperature t_w^h (°C)	22	20.34	19	24
Cooling set point temperature t_w^c (°C)	22	23.93	19	24

Table 7
Comparison of ATC and AAO solutions

Decision variables	ATC solution	AAO solution
Total objective function value (scaled)	0.013	0.098
Top-level design variables		
Thermal performance of office P_o (%)	1	0.81
Area of office A_o (m^2)	225.36	210.96
Thermal performance of workshop P_w (%)	0.94	0.834
Area of workshop A_w (m^2)	138	142.5
Linking variable y_s (m)	14	13.2
Subsystem 1 (office model): Variables		
Partial length of office y_o (m)	4.31	4.38
Window height g_o (m)	0.84	1.01
Mean zone velocity v_o (m/s)	0.14	0.1
Heating set point temperature t_o^h (°C)	21.21	21.75
Cooling set point temperature t_o^c (°C)	21.34	21.67
Subsystem 1 (workshop model): Variables		
Width of workshop x_w (m)	10.1	10.8
Window wall percentage g_w (%)	0.67	0.83
Mean zone velocity v_w (m/s)	0.18	0.20
Heating set point temperature t_w^h (°C)	20.34	21.56
Cooling set point temperature t_w^c (°C)	23.93	21.88

Table 8
Nomenclature for the design of thermal performance target

Variables/responses	Description	Units
T_{A}	target area of the three zone building	
T_{P}	target performance of the three zone building	
A_{w}	area of the workshop	
A_{o}	area of the office	
P_{w}	total annual energy used by the office	
P_{o}	thermal comfort of the office	
y_{s}	length of wall between the office and workshop	m
x_{w}	width of workshop's south wall	m
y_{o}	length of part of office wall	m
g_{w}	window height of workshop	m
g_{o}	window wall percentage of office	%
$t_{\mathrm{w}}^{\mathrm{h}}$	heating setpoint temperature of workshop	°C
$t_{\mathrm{o}}^{\mathrm{h}}$	heating setpoint temperature of office	°C
$t_{\mathrm{w}}^{\mathrm{c}}$	cooling setpoint temperature of workshop	°C
$t_{\mathrm{o}}^{\mathrm{c}}$	cooling setpoint temperature of office	°C
v_{w}	mean air velocity in workshop	m/s
v_{o}	mean air velocity in office	m/s
H_i^{w}	heating energy consumption in workshop at timestep i	J
C_i^{w}	cooling energy consumption in workshop at timestep i	J
H_i^{o}	heating energy consumption in workshop at timestep i	J
C_i^{o}	cooling energy consumption in workshop at timestep i	J
E_{refw}	maximum allowed total energy load of workshops	J
E_{refo}	maximum allowed total energy load of workshops	J
φ_i^{o}	thermal comfort in office at timestep i	PMV
φ_{w}''	maximum thermal comfort in workshop over all timesteps	PMV
φ_{o}''	maximum thermal comfort in office over all timesteps	PMV
φ_{ref}	maximum allowed thermal comfort	PMV
t_{w}''	maximum mean air temperature in workshops	°C
t_{w}'	minimum mean air temperature in workshops	°C
t_{o}''	maximum mean air temperature in office	°C
t_{w}'	minimum mean air temperature in office	°C
t_{ref}''	maximum allowed mean air temperature	°C
t_{ref}'	minimum allowed mean air temperature	°C

Other symbols	Description
$\mathbf{Z}$	overall building performance model
$\mathbf{Z}_{\mathrm{s1}}$	office decision model
$\mathbf{Z}_{\mathrm{s2}}$	workshop decision model
$\mathbf{R}_{\mathrm{Z}}$	top-level responses
$\mathbf{R}_{\mathrm{Zs}}$	subsystem responses
$\mathbf{y}_{\mathrm{Zs}}$	subsystem linking variables
$\tilde{\mathbf{x}}_{\mathrm{w}}$	local variables of workshop model
$\tilde{\mathbf{x}}_{\mathrm{o}}$	local variables of office model
ε_{R}	vector of tolerance variables for deviation between subsystem targets set at the top level and the responses computed by the subsystem level
ε_{y}	vector of tolerance variables for deviation between coordinating linking variables set at the top level and the linking variables computed by the subsystem level
k_{w}	scaling factor for area of workshops
k_{o}	scaling factor for area of office

5. Conclusions

Results from the pilot study demonstrate the potential of the ATC process for lending clarity and tractability to the typically complex decision-making problems in building performance analysis. In performance-based decision making, it is particularly beneficial to be able to determine compatible performance targets at all decision nodes on the basis of some overall specifications. Furthermore, at the end of the target cascading process, it is possible to systematically revisit the problem if some targets are not met, or if trade-offs between different performance goals do not appeal to the decision maker.

In addition, the decomposition approach allows the individuality of a local analysis task to be preserved, and so: (a) analysis tools can be invoked at particular levels and for specific needs avoiding information overload, (b) the decision-making space for a local problem is clearer since it only includes locally relevant variables and functional relationships, and (c) appropriate optimization algorithms can be invoked depending specifically on the formulation of the local analysis problem.

Several challenges were faced during this study. An example is identifying the typical decision variables that can be considered during the analysis process such as the external wall and roof assemblies. These are important decision variables in an exterior load dominated zone. However, they are difficult to include as decision variables due to non-smooth functional dependencies. Another challenge is identifying analysis tools that are relatively fast and also sensitive to changes in the values of decision variables. Finally, the solution obtained from the target cascading process depends on how the performance targets are weighted at the top level, and this poses problems typical of multi-criteria formulations.

Work in progress addresses these issues. Although we present its applicability through a pilot study, this strategy is valid for a broad class of complex building performance problems where the multiplicity of functions and performance specifications make it particularly difficult to retain the integrity of design decisions. Such applications also constitute work in progress (Table 8).

References

[1] C. Audet, J.E. Dennis Jr., A pattern Search Filter Method for Nonlinear programming without Derivatives. Houston, Texas: Rice University, Department of Computational and Applied Mathematics, 2000. Technical Report TR00-09.

[2] A.M. Brandt (Ed.), Foundations of optimum design in civil engineering, M. Nijhoff Publishers, Norwell, MA, 1989.

[3] L.G. Caldas, L.K. Norford, A design optimization tool based on a genetic algorithm, Automation in Construction 11 (2002) 173–184.

[4] R. Choudhary, A hierarchical optimization framework for simulation-based architectural design, PhD thesis, University of Michigan, 2004.

[5] R. Choudhary, A. Malkawi, Integration of CFD and genetic algorithms, Proceedings of the Eighth International Conference on Air Distribution in Rooms, Copenhagen, Denmark, 2002.

[6] R. Choudhary, A. Malkawi, P.Y. Papalambros, A hierarchical design optimization framework for building performance analysis, Proceedings of the Eighth International IBPSA Conference. Eindhoven, The Netherlands, 2003.

[7] D.A. Coley, S. Schukat, Low-energy design: combining computer-based optimization and human judgment, Building and Environment 37 (12) (2002) 1241–1247.

[8] D.B. Crawley, et al., Energyplus: energy simulation program, ASHRAE Journal 42 (4) (2000) 49–56.

[9] J.S. Gero, D.C. Neville, A.D. Radford, Energy in context: a multicriteria model for building design, Building and Environment 18 (3) (1983) 99–107.

[10] R. Hooke, T.A. Jeeves, Direct search solution of numerical and statistical problems, Journal of the Association for Computing Machinery 8 (2) (1961) 212–229.

[11] H. Jedrzejuk, W. Marks, Optimization of shape and functional structure of buildings as well as heat source utilization example, Building and Environment 37 (12) (2002) 1249–1253.

[12] D.R. Jones, C.D. Perttunen, B.E. Stuckman, Lipschitzian optimization without the Lipschitzian constant, Journal of Optimization Theory and Applications 79 (1) (1993) 157–181.

[13] H.M. Kim, et al., Analytical target cascading in automotive vehicle design, Proceedings of the 2001 ASME Design Automation Conference. Pittsburgh, Pennsylvania, September 9–12, 2001.

[14] H.M. Kim, Target cascading in optimal system design, PhD dissertation, University of Michigan, 2001.

[15] M. Kokkolaras, et al., Extension of target cascading formulation to the design of product families, Journal of Structural and Multi-Disciplinary Optimization 24 (4) (2002) 293–301.

[16] P.H. Levin, Use of graphs to decide the optimum layout of buildings, Architect 14 (1964) 809–815.

[17] R.S. Liggett, W.J. Mitchell, Optimal space planning in practice, Computer-Aided Design 13 (5) (1981) 277–288.

[18] A. Malkawi, Building energy design and optimization: intelligent computer-aided thermal design, PhD dissertation, Georgia Institute of Technology, 1994.

[19] J. Michalek, R. Choudhary, P.Y. Papalambros, Architectural layout design optimization, Engineering Optimization 34 (5) (2002) 461–484.

[20] N. Michelena, H. Park, P.Y Papalambros, Convergence properties of analytic target cascading, AIAA Journal 41 (5) (2003 (May)) 897–901.

[21] N. Michelena, H.M. Kim, P.Y. Papalambros, A system partitioning and optimization approach to target cascading, in: U. Lindemann, et al., (Eds.), Proceedings of 12th Int. Conf. on Engineering Design, vol. 2, 1999, pp. 1109–1112 Munich.

[22] M. Monks, B.M. Oh, D. Julie, Audioptimization: goal-based acoustic design, IEEE Journal 0272–1716 (2000) 76–91.

[23] P.Y. Papalambros, Analytical target cascading in product development, Proceedings of the 3rd ASMO UK/ISSMO Conference on Engineering Design Optimization. Harrogate, North Yorkshire, England, July 9–10, 2001.

[24] P.Y. Papalambros, The optimization paradigm in engineering design: promises and challenges, Computer-Aided Design 34 (12) (2002) 939–951.

[25] C. Park, G. Augenbroe, T. Messadi, Daylighting optimization in smart façade systems, Proceedings of the Eighth International IBPSA Conference, The Netherlands, 2003.

[26] K. Peippo, P.D. Lund, E. Vartiainen, Multivariate optimization of design trade-offs for solar low energy buildings, Energy and Buildings 29 (2) (1999) 189–205.

[27] J. Pohl, L. Myers, A. Chapman, ICADS: An intelligent computer-aided design environment, Ashrae Transactions 96 (2) (1990) 473–480.

[28] A.D. Radford, J.S. Gero, Design by optimization in architecture and building, Van Nostrand Reinhold, New York, 1988.

[29] A. Saporito, et al., Multi-parameter building thermal analysis using the lattice method for global optimization, Energy and Buildings 33 (3) (2001) 267–274.

[30] M.J. Sasena, Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations, PhD dissertation, University of Michigan, 2002.

[31] E. Shaviv, Y.E. Kalay, Combined procedural and heuristic method to energy-conscious building design and evaluation, in: Y.E. Kalay (Ed.), Evaluating and predicting design performance, John Wiley and Sons, New York, 1992.

[32] I.H. Sloan, S. Joe, Lattice methods for multiple integration, Oxford University Press, New York, 1994.

[33] T. Wagner, A general decomposition methodology for optimal system design, PhD dissertation, University of Michigan, 1993.

[34] M. Wetter, E. Polak, A convergent optimization method using pattern search algorithms with adaptive precision simulation, Proceedings of the Eighth International IBPSA Conference. Eindhoven, The Netherlands, 2003.

[35] M. Wetter, J. Wright, Comparison of a generalized pattern search and a genetic algorithm optimization method, Proceedings of the Eighth International IBPSA Conference. Eindhoven, The Netherlands, 2003.

[36] A.J. Wilson, A.B. Templeman, An approach to the optimum thermal design of office buildings, Building and Environment 11 (1) (1976) 39–40.

[37] J.A. Wright, H.A. Loosemore, R. Farmani, Optimization of building thermal design and control by multi-criterion genetic algorithm, Energy and Buildings 34 (9) (2002) 959–972.

[38] M. Zaheer-uddin, G.R. Zheng, Optimal control of time-scheduled heating, ventilating and air conditioning processes in buildings, Energy Conversion and Management 41 (1) (2000) 49–60.