A METHOD FOR THE DESIGN AND ANALYSIS OF PARAMETRIC BUILDING ENERGY MODELS

Kevin B. Pratt¹ and David E. Bosworth²

¹Department of Architecture, Cornell University, Ithaca, New York, USA
²Program of Computer Graphics, Cornell University, Ithaca, New York, USA

ABSTRACT
This paper presents new methods for understanding specific building performance trends and their dependencies by parameterizing digital building models over multiple variables in order to create a high dimensional design space that can be rapidly simulated, analyzed, and visualized. These methods allow for the potential delineation of multiple areas of the design space that are characterized by relatively high performance instead of a single optimum. The ability to identify clusters of high performance is particularly helpful to architectural design processes, which must take into consideration variables that cannot be parameterized and resist quantitative methods of optimization.

INTRODUCTION
In 1973 Horst Rittel and Melvin Webber proposed classifying certain design problems as wicked problems, which are defined as problems that are difficult to solve because their solution depends on resolving relationships between a host of often contradictory qualitative and quantitative interdependent factors. Wicked problems are characterized as having no optimal, but rather only better or worse solutions. Such problems tend to be unique, which precludes the effective application of prescriptive or historically determined methods to solve them. Environmentally coupled energy efficient building design falls into this category of problem, and design processes currently used in the creation of energy efficient buildings are often ill-suited to address the challenges such problems pose.

Why is this so? The answer seems to lie in the fact that the design space, which we define as the implicit range of potential variation or list of potential options (i.e. the potential solution set) being studied in relationship to a particular design problem, as it is typically constructed in the earliest stages of sustainable design processes, is constrained, intermittently populated and information-poor. Rittel observes that the set of potential solutions to a wicked problem is dependent on a concrete definition of such a space, that in fact the definition of the space is in a real sense the solution itself. We believe that is reasonable to assume that poorly articulated design spaces will result in less than optimal, or less good, solutions.

Current sustainable design processes can often be broken into two phases. Typically, in the first phase the designer (usually an architect) develops various combinations of form, material, and siting with the goal of producing a preliminary design that achieves low combined loads (heating, cooling, ventilation & lighting), while simultaneously satisfying other aesthetic, spatial, programmatic and social imperatives. The systems, both architectural and mechanical, that meet these loads are then specified and optimized in the second phase, usually by an engineer or trade subcontractor. The design space is constructed in the first, and therefore most critical phase, thereby proscribing much of the work that follows.

This design space is constrained by the fact that commonly used computer aided design tools produce static digital representations of particular building proposals and it is difficult, given time and economic constraints, to produce a wide range of highly variable design options. The design space is typically intermittently populated because the variations produced by architects using these tools tend to be discrete rather than continuous, thus omitting “in-between solutions”. This approach is exemplified by the typological matrices often found in textbooks, which describe the differences between bar buildings oriented east west, bar buildings oriented north south, L shaped buildings, cruciform buildings and other easily recognizable building plan configurations (Stein 2006).

In recent years, the design professions have begun experimenting with parametric design tools, such as Grasshopper (McNeel 2010), which can generate a wide range of variation within a defined parameter space (solving the intermittent population problem), and generative algorithms, which use feedback loops to escape a priori limitations on the scope of that parameter space (Schumacher 2010). However, such un- or weakly constrained and well-populated design spaces are still information-poor. This information poverty becomes a critical shortcoming as it is difficult to draw distinctions between the very large number of potential design options produced.
In general, we define information poverty as a lack of predictive information which would give insight into how a design might perform in the “real world”. Performance, when measured terms of ecosystem impact, energy use, and other quantifiable metrics, is of critical importance in sustainable design, while performance, measured in light of traditional architectural values such as “firmness, comity and delight” (Vitruvius, 1st century CE) remains important within the larger realm of architectural discourse. Architectural methods of assessing the latter have been greatly enhanced in recent years by computational methods such as photorealistic rendering. However, designers still tend to rely on rules-of-thumb that offer generic “environmental” design advice based on coarse measures of local conditions like degree-day and latitude to assess the former. This approach was pioneered by the Olgyay brothers in the nineteen fifties (Olgyay 1963) and is currently embedded in both prescriptive energy codes like ASHRAE 189.1 and green rating systems like Leadership in Energy and Environmental Design (LEED) (USGBC 2010). Given that sustainable design problems tend to be wicked, and thus distinctly novel and unique, this approach is often inadequate for the purposes of constructing design spaces that address sustainability concerns.

Energy simulation, which can be used to develop predictive information that would address the information poverty issue, is routinely used in the second phase of sustainable design processes, to calculate loads, verify that performance criteria have been met, optimize mechanical systems, and document energy code compliance (Ellis et al 2006). It is rarely used in the first phase because: a) current simulation software makes it difficult to efficiently explore design options, b) visualization methods are not well suited for the evaluation of simulation results in relation to non-performative or qualitative design goals and c) The architectural profession lacks established methodologies that incorporate performative analyses into the development of early stage designs.

The goal of the work described in this paper is to develop new tools and methodologies that allow designers to rapidly create, simulate, and understand expansive, well-populated, information-rich design spaces in the early stages of design. To this end, we propose marrying parametric methods for the generation of form with high throughput methods of energy simulation and novel forms of data analysis and visualization.

High dimensional design spaces of the kind we describe herein, which have more than three parametric variables and thus more than three axes of variation, are particularly difficult to simulate, analyse and visualize. Such spaces often contain more than 50,000 discrete models, far more than can be effectively simulated sequentially; we apply simultaneous parallel simulation to manage what would otherwise be unacceptable long simulation times. Running so many simulations creates inordinately large data sets, thus statistical methods are critical to gain insight into design space topology. Finally, because humans generally have difficulty understanding more than three spatial dimensions we are precluded from completely visualizing a totality of the results as points in Cartesian space. Therefore graphical methods for exploring the data, viewing lower dimensional cuts through high dimensional design spaces, and for viewing the totality of the design space in abstract form are proposed.

PREVIOUS WORK
Ideally, energy efficient buildings take advantage of available resources by coupling thermodynamically to the natural energy sinks and sources in the surrounding environment. This requires that building geometry, orientation, mass, and materials are proportioned correctly to take advantage of local climate dynamics. In the past, tuning a building to a site has been done through rule of thumb methods. The availability of digital simulation has allowed for the specificity of a single building design to be interrogated, which had led to the widespread use of simulation for purposes of evaluating code and rating system compliance, but little else. Attempts have been made to improve the generality of rule-of-thumb procedures by using bulk simulations to form matrices of optimized building designs with the intent of interpolating between instances for a specific site (Brown 2001). Other means of dealing with the specificity of the problem have been the use of search techniques using one of many single or multi-variable optimization algorithms. Any quantifiable design variable that is amenable to the application of a fitness function is relatively easily subjected to automated search and optimization using various methods such as hill climber, pareto-front, genetic algorithms, simulated annealing, particle swarm, and many others (Chela 2009, Lagois 2010, Tuhus-Dubrow 2010). Such optimization routines can be useful and efficient, but they fail to describe the entire design space. We believe that describing the entire design space is useful at the earliest stages of design because: a) The large dataset allows for a better analysis of variable dependency and coupling, b) It is often just as useful to know which variants fail as it is to know which ones are successful. This may lead, we hope, to a more intuitive sense of the topology of the design space. c) The set of solutions which may satisfy both parameterized and unparameterized criteria may in fact fall far from the quantitative optima.

DESIGN OF PARAMETERIZED MODELS
Of critical importance in creating a parameterized design space is an understanding of how to parameterize a problem in a way that produces
coherent and analyzable results. In practice this is a non-trivial issue, as real architectural problems often operate under a series of unique constraints that preclude the use of continuous change to describe the dynamic model. As a general rule, we categorize all variables as either continuous or discrete. Continuous variables are those in which the variation can be described graphically as a continuous curve, such as the variation of percentage of glazing on a façade. Discrete variables are those which cannot be so described, wherein the variable consists of a list of specific design options. A good example of a discrete variable parameterization would be providing the simulation engine a series of unique floor plates to be tested. Practically, it is difficult to discern variable dependencies when one or more of the parameters are discrete. Furthermore, the use of discrete variables renders the local derivatives of the search space topology less meaningful. Thus, the use of such parameters is best avoided.

It is also worth noting that there are parameters that are not easily described mathematically, but whose variation can be qualitatively assessed by examining the model variations. This is especially true of certain architectural qualities, such as spatial proportion, lighting effect, or material effect; thus it is important to construct the parameterized model in such a way that these qualities are easily discerned through a visual examination of the digital model. This is also why it is critical to be able to concurrently visualize both the simulation data output and the model data input when using such tools in an architectural design process.

MODEL AND PARAMETER CREATION

One of the most significant barriers keeping architects and designers from parametrically probing design options in relation to energy performance is the discrepancy between the CAD tools used for creating building geometry and the software packages used for running simulations and viewing the results. To resolve this problem a suite of tools called sustainParametrics and exportZoned were created as Ruby plugins for Google SketchUp (Google 2011). These tools aid in the development of the required geometry and meta-data required to run a parametric simulation. Very few constraints are placed on the methods and processes used to build the model. It is only required that separate thermal zones be separated onto different layers that conform to a naming convention. Materials are assigned to the surfaces through the materials palette native to SketchUp. The exportZoned tools translate the SketchUp geometry to a model compliant with simulation environment modelling protocols. Translation of the building geometry to a building model capable of simulation by EnergyPlus (EERE 2011) requires that adjacent thermal surfaces are identified, that adjacent surfaces are the same size, that materials and constructions are assigned appropriately, and that the child-parent relationship of sub-surfaces to parent surfaces is maintained. These manipulations are performed during the import stage to the simulation manager. Parametric control of the building is accomplished through a concept borrowed from digital animation known as character or skeletal rigging. After a seed building is modelled, using any processes or methods the architect wishes, pertinent surfaces, edges, and groups of objects are associated with a hierarchical set of controls, which are analogous to the skeleton used in character rigging. For continuous geometric variables geometric transformation matrices are employed, which allow for translation, rotation, shearing, mirroring, and scaling of building components. For discrete variables array lists of inputs are required. A graphical user interface (GUI) based on the wxSU platform (Ellis 2008) grants dynamic access to modification of the SketchUp model by sweeping through all of the building transformations (Figure 1). Dynamic animation of the building transformations provides quick graphical troubleshooting of the parameterization of the building. SustainParametrics creates the full enumeration of the variations; from every combination of the parameters a building model for simulation is created. The models are simulated using EnergyPlus to produce a general yearly building energy use metric. It is also possible to produce a carbon equivalent (CO₂e) metric to assess global warming impacts.

SIMULATION AND BATCHING

All of the simulations were run using Sustain, a user interface developed at the Program in Computer Graphics at Cornell University (PCG). Sustain performs simulation management on local computers or on a cluster, and management and display of the
simulation results. Large simulation runs are performed on the PCG cluster, an Intel based homogeneous cluster of 256 cores. Sustain supports EnergyPlus and Radiance (LBNL 2011).

DATA VISUALIZATION
Multiscalar visualization techniques were developed that promote an understanding of the effects of design parameters on energy performance. The high dimensional design space topology is particularly difficult to visualize and even discuss. Nominally, four input variables and one dependent output variable can be visualized in Cartesian space, using X, Y and Z co-ordinates for three, colour for the fourth, and time (using animation) for the fifth. Where more variables are present one can describe the characteristics of the topology as being similar to those of a three dimensional surface that results from a two dimensional optimization function where the dependent variable is expressed by the Z co-ordinate.

Speaking analogically, we wish to understand the data “landscape” in order to discern very sharp peaks, where small displacements have drastic effects on the results, from broad plateaus, where large changes have minor effects on the results. Decisions in the low slope areas of the design space are less critical than decisions in the high slope areas. We also want to discern “ridges” where changes in one parameter are critical to building performance while changes in another are not. Finally, we want to find “valleys” or areas of the space that hold little promise and can be rapidly abandoned.

Visualizing the entire design space is important because it allows designers to differentiate critical decisions from those where a broad range of solutions will suffice. Optimization strategies, like genetic algorithms or pareto-front methods, essentially find pathways through the design space that can be thought of as analogous to lines drawn on a three dimensional surface. While these pathways can be analysed, they do not allow for a holistic understanding of the topology, which in turn can allow a designer to key in on aspects of the building that have a strong effect on energy use, while allowing other aspects of the building, which may have weaker effects on energy use, to be determined by other design criteria like views, circulation, program, and aesthetics.

Graphical Interface
The results from a parameterized simulation run are displayed in three windows that are visible simultaneously and update interactively: a detailed view of one iteration of the building in the Sustain main console, a three-dimensional voxel-plot of the simulation results displaying three axes of parameterization at once, and multi-axis batch controller for data navigation. The batch controller displays all of the axes of parameterization at once as navigable sliders and transposes a colour coded parallel co-ordinate plot to identify particular buildings within the search space (Figure 2).

Main Console
The main console of Sustain displays the building and its environment and allows the modeller to probe the model’s assembly details, location, and associated weather file. Detailed simulation using EnergyPlus is managed from the main console and results of the simulations are viewed.

Voxel-Plot
Simulation results from three of the parameters are viewed as cuts through the search space using a colorized three dimensional interactive voxel-plot. Each voxel displayed represents one combination of all of the parameters: one iteration of the building. The parameters not viewed in the voxel space are set to a default value, which is adjustable via a slider in the batch controller. The results of the simulations are colour-mapped to the voxels in either a global scale, where the colour gradation ranges from the lowest to highest in the entire simulation run, or to a local scale, where the colour gradation ranges from the lowest to the highest within the subset of the combinations displayed. The main console and the voxel-plot display are linked so that selection of a single voxel displayed in the search space loads the associated building in the main console, facilitating simultaneous understanding of the specific building and its location within the space of all possibilities.

**Batch Controller**

The batch controller provides navigation control of the voxel space and simultaneously allows for a broader visualization of the entire search space, as it is both controller and display. Sliders and buttons which control the display of the search space by adjusting the slice through the full data set are on the bottom and left side of the central display. The buttons on the left control which variables are displayed and which are held constant, the buttons and slider at the bottom delineate the subset of the search space visible in the voxel-plot. At the centre of the window each parameter is displayed on a horizontal axis as an interactive slider; these axes are then stacked and a highlighted line is drawn through the specific parameter values that describes the model visible in the main console. The remaining buildings in the sub-set are drawn in the background as lines, colour coded to match the voxel-plot, forming a parallel co-ordinate plot (Fua et al 1999). The subset of the search space is defined as all the buildings that fall within the space described by the two fixed and three free variables and also fall either above or below a performance cut-off, defined as a percentage of the entire data set. This cut-off point is controlled by the user with an interactive slider. The colour mapping of the voxel space is used for continuity and direct comparison, and can be switched from a global to local scale, thus is possible to see the sub set in relation to both the whole data set and only the subset on display. There are also buttons which reset the display to the highest and lowest performing regions in the search space; this is necessary as it can be difficult to find the edges of the search space through direct manipulation of the batch controller, as the totality of the multi-dimensional data set is never entirely visible in the voxel-plot.

The parallel co-ordinate plot formed by stacking the parameter axes and drawing color coded lines to describe the relevant building variations enables the most global view of the search space, while the voxel-plot gives insight about the local topology. When used together they can be remarkably effective tools for interpreting the simulation results. When coupled to the model view main console, the system creates a multi-level layering of both qualitative and quantitative information that allows for a holistic understanding of the design experiment.

**EXAMPLE SIMULATION**

As a proof of concept we parameterized a design for a simple passive solar house on a sloping site (figure 3) The three zone model is varied in five dimensions:

- **Degree of ground coupling** in 7 steps which progressively increase below grade wall area
- **Orientation** in 13 discrete fifteen degree increments of rotation about solar South
- **Amount of glazing parallel to primary axis** in 9 steps from 6% to 56%
- **Scale of horizontal shading devices** in 7 steps from no shading to complete shading
- **Amount of internal thermal mass** in 6 steps of increasing floor slab thickness from 2 inches to 12 inches of concrete

The remainder of the building geometry was assigned assemblies typical of residential construction in upstate New York. The resultant search space contained 34,398 discrete models; these were simulated using EnergyPlus via Sustain in less than two hours. Each model was subjected to a full year analysis using Binghamton, NY, USA as the reference climate and purchased air as an ideal load air system. Given the potentially enormous size of the resultant data set (depending on the number of variables requested Energy Plus can generate hundreds of megabytes of output data for a single simulation), a single output metric was used for model evaluation: carbon footprint by unit area (kg/CO₂e/m²). The data was written to a custom binary file format developed for Sustain which allows for rapid data access.

The five dimensional design space was visualized using dynamic cuts through the data set using the

![Figure 3. Example Model Parameterization](image)
Figure 4. a) Single Best Case Building & Voxel-Plot, b) Single Worst Case Building & Voxel-Plot, c) Top Performing 15% Voxel-Plot, d) Worst Performing 15% Voxel-Plot

Figure 5. Top Performing 15% Batch Controller with Superimposed Linegraph Plot

Figure 6. Top 15% from Shaded Model

Figure 7. Top 15% from Shaded Model
visualization methods described above. Obviously, the small passive solar house, with its familiar section, was modelled precisely because good assumptions pertaining to its behaviour can be made and thus the validity of the results quickly evaluated. The utility of the voxel-plot is readily apparent. We can find the highest performing variant (figure 4a) which, unsurprisingly, corresponds to a partly shaded south facing variant with ample (but not overmuch) glazing, lots of thermal mass, and strong ground coupling. We can also look for the worst variant (figure 4b): low mass, too much glazing facing due west; again as expected. We can also begin to search for families of high performing variants. We can fix two variables (mass and glazing area) and make a voxel plot of the top performing fifteen percent of variants (figure 4c). We can do the same for the worst performing fifteen percent (figure 4d) and see the two lobe distribution with clusters at extreme east west orientations that share a minimal ground couple.

Looking at the batch controller parallel co-ordinate plot displaying the top fifteen per cent performance range (figure 5) one can intuit further information about the design space. The concentration of blue lines at the left of the ground level axis indicates that high performing options are almost always strongly ground coupled. The distribution of the coloured lines across the orientation axis indicates that there are good performing options at all orientations; the clustering of blue lines in the centre of the axis suggest that the options that face south perform better than those which face east or west. The distribution of lines across the overhang axis, where blue lines are primarily distributed on the left side of the graph (smaller overhangs) suggests that less overhang is actually better; this is probably due to the fact that the reference climate used is heat load dominated and summer skies are often overcast. The concentration of blue lines at the right of the floor thickness axis indicates that greater thermal mass leads to better performance (again probably a consequence of using a heat load dominated climate). The distribution of lines across the window size axis suggest that median values are optimal but good options occur across the range of parameterization, since we are, after all, looking at a subset that includes the top fifteen percent of all performance values in the entire search space.

By stepping back and looking at the general shape of the batch controller graph we draw further conclusions: that performance is strongly coupled to ground coupling and available thermal mass, somewhat coupled to overhang length, and more weakly coupled to window size. We can also see, by looking at the zig-zag shape of the blue lines, that the relative best performance will be achieved by maximizing thermal mass, using a minimal shading overhang, orienting the building towards the south, and burying the north wall.

As a further test of the method we ran a second design space simulation, identical to the first in all respects except for the fact that we placed a shading obstruction (nominally a tree) south of the building. When we look at the resultant voxel-plot and batch controller graphs (figures 6 & 7), we see that the regions of optimal performance shift away from the south facing buildings towards those that face either east or west; one also sees that buildings with large windows do not perform as well. This is to be expected given that the original design depends heavily on passive solar gains in the winter to offset thermal losses; this mechanism is disrupted by preventing sunlight from reaching the south façade of the building. This experiment also shows a simple example of a multi-lobe distribution of optima within the design space, where regions of high performance are non-contiguous. Practically, this means that there is more than one sufficing solution to the quantitative aspects of the design problem. As we simulate more complex buildings on more complex sites, we expect to see more complex multi-lobe performance distributions.

FURTHER WORK

Parameterizing complex buildings

There is a clear need to parameterize and simulate more complex buildings. As an initial experiment we have created a parameterized model based on the design of Milstein Hall, the new home of the Department of Architecture at Cornell University.

![Figure 8. Milstein Hall Parameterized Model](image)

The model slides between the adjacent buildings as fenestration and skylight percentages vary (figure 8). It also fills in the large undercroft space in the original building with conditioned space in an attempt assess the effect of the cantilever of the upper story volume. Initial results suggest that the simulated design space (> 60,000 models) is complex and that this data set will benefit from further statistical analysis.

Conducting further statistical analysis.

The Batch Visualizer is currently capable of calculating main and secondary effects (Box 2005)
but we have yet to find an effective method of displaying this information in a graphical interface when inspecting a high order (i.e. > three variables) search space. We believe that this information will prove useful as it is a direct quantification of variable sensitivity and interdependency. One challenge that we face lies in the fact that traditional mathematical means of conveying this information, in the form of tables, graphs, etc., are of little use to designers who tend to lack the technical background necessary to make use of them.

CONCLUSIONS
A prototype method for creating, simulating, analysing, and visualizing expansive, well-populated, information-rich design spaces has been presented. The parameterization of selected variables within a digital model to construct the entire design space is effective, and has not been done previously due to the computational expense associated with performing so many discrete simulations. Visualizing the high dimensional design space that results remains a difficult unsolved problem, and we suspect that successful methods will be specific to specific disciplines. Our experiments have found that in the realm of building design the simultaneous display of data in relation to geometry as well as in abstract form helps foster an understanding of the entire design space, which in turn helps designers find “better” solutions to the wicked problem of creating sustainable buildings.

Although the initial example shown here is of limited scope, the method shows promise and will, if further developed, prove useful in the design of sustainable buildings. In a sense the method is, conceptually, an anti-optimization strategy. It allows the designer to retain full control of decision making within the design space, while at the same time making evident the performative consequences of particular decisions. Thus it seeks to persuade rather than dictate, and it offers options rather than specific answers. For this reason, we believe it will be attractive to designers, especially those who see sustainability as an externally applied set of constraints that inhibits their creative processes.

ACKNOWLEDGEMENTS
This material is based upon work supported by the Department of Energy under Award Number DE-EE0003921. Tests were run on prototype software created with the assistance of Lars Schumann and Nathaniel Jones.

REFERENCES
DOE2 (2010) “eQuest, the quick energy simulation tool”, http://doe2.com/equest/
Google (2011) "Google SketchUp", http://sketchup.google.com: