QUALITY ASSURANCE - SIMULATION AND THE REAL WORLD

Michael Donn
Centre for Building Performance Research (CBPR)
School of Architecture, Victoria University
Wellington 6001,New Zealand
michael.donn@vuw.ac.nz

ABSTRACT

This research asks the nature of environmental design information used by building designers. The goal is to identify commonality in the types of information that design tools should produce. It presumes improving use of design tools will lead to improved building performance. Through practitioner interviews, it investigates application of design decision support tools by building designers. It proposes a means of increasing designers use of these tools. This proposal derives from observation that systematic Quality Assurance (QA) systems are seldom used with simulation-based tools. The proposal is a QA system comprising a) a simulation veracity test akin to the Turing Test of computer intelligence; b) an Internet database of building performance information; c) post analysis tools that define the reliability of design tool output.

INTRODUCTION

This research[1] asks: *what is the nature of the environmental design information sought by building designers?* It assumes that improved building performance is the goal of all designers. For the research, we interviewed users of building environmental design decision support tools. The goal was to identify whether common needs exist across different environmental design topics for particular types of tools and information.

All environmental design decision support tools are "simulations" of some imagined reality. These simulations can be charts of building performance versus window size in a solar house design handbook; an R-value calculation; a wind tunnel test; or use of a computer to predict the performance of a building. Complex computer programs like DOE2[2] and RADIANCE[3] that are conventionally known as 'simulation' programs, are just more detailed, and potentially more realistic simulation tools.

THE BACKGROUND

The premise is that architects need the environmental analysis skills to adapt the architectural *precedents*[4] they are encouraged to study. Using these precedent buildings in new locations requires more than a mere assertion that architecture has *profound significance* or *embodies timeless laws*[5]. What they require is a systematic understanding of the relationship of those precedents to the environment **-** a numerate designer.

The research turns always on this one critical point: to what degree does the role of the architect in the building design team require analytical skills? The conventional notion[6] is of the architect as team leader. Building science researchers frequently conclude that a design approach compatible with these conventions[7] that is also to deliver quality environments should place central importance on *early design decision making*. Many spend long hours developing tools that they intend will improve the effectiveness of early design decisions [8].

THE "EXPERIMENT"

For the research project we interviewed practitioners on use of a range of different design decision support tools. The interviews varied in delivery and question. Forty-six of 130 New Zealand architects and designers in workshops introducing concepts of passive solar house design returned a follow-up questionnaire. We interviewed architects in sixteen of the twenty-three firms involved in wind tunnel tests as part of the planning permission process for 51 high rise buildings in Wellington. We also interviewed two architect clients of CBPR thermal and visual simulation services. Eighty practitioners in New Zealand responded to telephone or personal interviews about their use of energy efficiency design decision support tools. Twenty users of computer thermal simulation programs in the USA responded to 'phone interviews.

The interviews showed all these different modes of simulation have problems that reduce the clarity of the relationship between prediction and reality. These problems fall into the following classes:

- model preparation time limits;
- no clear guidance as to the important features of a building that should be modelled well, and the features whose effect on predicted performance is insignificant;
- minimal quality control systems that allow the simulation user to ensure the relevance and accuracy of their recommendations;
- lack of performance guidelines for buildings that provide a basis for understanding the recommendations from the simulation;
- lack of tools for summarising and detecting patterns within the information overload that well-applied simulation can produce when exploring design scenarios.

Limited preparation time

Half the passive solar house design seminars were spent in lectures, the other half in workshops. Participants showed strong interest in learning the calculation techniques and the Rules of Thumb. When trying to prove to themselves that their designs would ‘work’, they used the calculations. This is in complete contradiction to the findings of the general survey of New Zealand practitioners’ use of energy efficiency design tools. Their call was for tools that they could use with little effort. No client would apparently pay for the extra analysis time.

USA simulation program users’ responses agree with researchers’ conclusions[9] that the most important phase of the design process is the first, so-called conceptual design phase. Participants in the New Zealand and USA surveys did not comment directly on the time taken to use a simulation program. However, when asked to suggest improvements to simulation software many of the USA users said they wanted easier to use programs. The principal benefit was to reduce simulation time.

A very high proportion (79%) of the USA participants saw a high value in continued and expanded use of simulation in building design. None commented that simulation was better applied later in design.

Concerning “*practical difficulties ... in carrying out wind tunnel tests yourself* “, typical comments described the time this type of simulation takes away from the design process: “*Not enough people in the office to spare someone for that time. Not confident to have done it efficiently and to have come out with a good report. … Is it necessary? Why can’t others do it? ... Design time is usually quite short, and anything adding to that is an obstacle.*” ... Another aspect of time is the cost of doing a wind tunnel simulation. Comments on this aspect included: “*Expensive, one* *__time__*, *models and analysis taken into account. ..*__*Time*__, *expertise, cost all have an effect* - *testing takes much longer than it should as a result of the lack of experience, wasting time, costing money while gaining experience, not economical or time efficient for developer”*. They argued for a simplification of the method. The implication was that simplification would produce quicker turnaround of applications in the Council approval process - a saving in time.

Time was the most significant aspect of any comments from architect-clients about CBPR thermal and visual simulations. The architects suggested that the extra time required to consider fully the options put forward by the CBPR would have put them behind schedule. They could not afford to do this as they risked losing money at current low design fee levels. “*[We] had limited funds, therefore we had to produce a design quickly to come within the fees we were being paid. [Because of the] minimum fee we weren’t interested in pursuing alternatives unless we were paid for it.”*

Lack of guidance on which features are significant

The participants in the passive solar house seminars tried the “simulation” formulae provided in the manual[10] almost at random to sort out what worked. They did not behave as if they had any idea, even after the lectures, which building features would have the greatest effect. They did not try to find connections between the “Rules of thumb” and the calculations. Rules of thumb typically specify what size a building feature (window size, wall thickness, thermal storage) “should” be but they do not normally specify why.

The solution seems to be to produce simulation programs that are flexible. Designers would use them to explore the interaction of all parameters in a design. This type of simulation requires more time and resources than current programs. The highest priority is for simulation output to include a ranked list of the building parameters that affect performance the most.

When asked to describe the priorities they would set in establishing courses for new users of simulation, the USA simulation program users talked of teaching 'scepticism'. They wanted to imbue a distrust of the Black Box. They were trying to convey the importance of knowing what to model. They were also saying the user of simulation must have sufficient curiosity to ask why each simulation produces its particular output.

A 2:1 majority of the participants in the USA survey saw a basic knowledge of what was important to simulate as an essential skill. Current mainstream computer-based building performance simulation tools do not provide guidance to help the user to learn which parameters are significant for the current simulation.

Architects interviewed about wind tunnel tests showed a singular lack of understanding of the concept of "accuracy" in the results. Comments were common about the "inaccuracy" of the modelling process. This lead to questioning of the conclusions and recommendations. A lack of detail in the model raised doubts about the credibility of the results of the tests. Abstracting of the model like this so that only the essentials necessary for simulation accuracy are described is a normal part of any modelling process. For example, a computer simulation tool for wind flow modelling used by these architects would have to include some very sophisticated "help" files and "verification" options. They would need these to check that in the interests of what they saw as accuracy, they had not made the model too complex. Such a procedures might appear as a 'verify' button in the program that reports a 1000-fold increase in calculation speed with an undetectable change in predicted flow if they remove all window mullions from the model used in the calculation.

RADIANCE like all physically based rendering engines has many input variables describing a scene. This gives the user an apparently infinite number of ways of `getting it wrong or right'. The complexity of the definitions typically compounds the uncertainty of the environment variables in the input. It also does not help that these variables have few readily apparent absolute `real world' correlates. They are artefacts of the approximations needed to complete the simulation.

Program developers aim much of their current work developing improved interfaces for building performance simulation programs at improving their accessibility. Designers already perceive pictures or renderings of the appearance of a space under a given illumination as 'user-friendly'. There is a high risk that many designers using these 'user-friendly' interfaces will be so seduced by the graphic output that critical examination of other outputs will be given a very low priority.

Lack of performance guidelines

In the solar seminars, the participants wanted to know whether the results of their calculations showed their building designs were "good". The only way in which this could be done in the seminars was through comparison with other groups' calculations for their buildings. This is the conventional approach. As performance simulation can never model all operational parameters, simulationists argue that two building models must be compared using the same performance calculation tool. They can then attribute differences in performance to differences in the two designs. Definition of the so-called 'standard' or 'base' building requires careful planning and documentation[11].

Any building performance simulation program must contain tools to assist such performance comparisons. If the tool is well designed, it will also contain a) a means for storing and comparing past simulations; b) tools for representing graphically the relationship between building performance and building design parameters; and c) a search mechanism for finding building performance data matching the type of building and occupancy being analysed.

The USA users of simulation programs described features of a 'user-friendly' program that are beyond the conventional image of a Graphic User Interface (GUI). Besides WIMPS[12] features, they described features that would help users to interpret the output data and to judge the building performance.

When asked whether they used *"particular design techniques to improve wind effects..."* the architects interviewed about building aerodynamics said things like *"No. Driven by economics"* and "*Don't have any rules of thumb*". However, during the interview most showed an awareness of several techniques. Overall, most replied that they found "*taking part in a wind tunnel test helped* [them] *in the design process*". However, the responses to the survey suggest that there is concern about the meaning of the wind tunnel measurements: *"Don't believe that the wind tunnel simulation is accurate enough for some sites to give a sensible solution to the problem."*

A design tool for building aerodynamics also requires an analytical "tool" which permits the user a) to specify what performance targets are to be met; and b) to request automated tests of compliance. Both the current wind tunnel test and any likely future CFD computer program require add-on "tools" of this type if architects are to use them.

Figure 1 Art Gallery, "reference" 150 lux spotlight defines "good" performance in the picture itself

The most challenging problem encountered with the use of rendered output arises in presentations of the light simulations to the client. Without a visual reference point in the picture, no amount of annotation or graphical overlay can convince them of the results. They are as aware as anyone that by adjusting the "exposure" of our "digital camera" we can alter pictures to make very bright conditions look innocuous, or very dark conditions, pleasant.

Expressing the results of performance analysis in a form and format that we can understand is a little understood science. Graphical representations where we represent the light intensities as a series of coloured contour lines across the rendered picture represent only quantity and not quality. Clients also find them hard to comprehend.

A specification of a design simulation tool for architects must include associated data analysis and manipulation "tools" that help comprehension of the numerical output. The spotlight "placed" in the rendered art gallery scene in **Figure 1** is just such a device. It provides a well-understood reference **-** 150 lux taken from the design brief as the target illuminance **-** against which to judge the natural light. It should not be up to the ingenious user to devise "tools" like this to calibrate simulation output. A simulation program that is a genuine design decision support tool should include the means to compare and contrast design options. It needs inbuilt indices of performance that describe the building in everyday language. The language of comparison should use real concepts like "*warm"* and "*cold"* not *24ºC* and *17°C; or* "*good for reading"* and *"adequate for egress"*, not *'400 lux'* and '20 lux'.

Lack of quality assurance tools in the simulation

We did not directly evaluate the influence of quality control on the accuracy or otherwise of the solar house designers' use of the *Design for the Sun* manual[13]. However the participants in the seminars required independent verification of the accuracy of their calculations. Often designers abandon calculations once they leave such specialist seminars. This is not because they see performance data as irrelevant, but because they have no independent verification of the calculation results.

A computer simulation program for solar house design must include an expert 'advisor' (an agent[14]?) to help with interpretation of the output. An advisor like this would be an essential aid for all users of simulation. Some simulationists already have their own version of such an expert advisor. They keep a library or database of old simulations. They look these up to discover which simulation they have done in the past is most like the current one. They then ensure the currently predicted performance is not to far from that in the database. A simply accessed, constantly updated database of performance benchmarks like this is essential to quality simulations.

Systems for documenting the process by which the simulationist has ensured that the building modelled is genuinely the building that is on the plans are essential. Without them, computer simulation remains a rare activity undertaken by a small group of aficionados and gurus - a priesthood if you will.

The architects interviewed about wind tunnel tests largely did not participate in wind tunnel tests themselves. Thus, we did not expect them to understand the detail of Quality Assurance tests in building aerodynamics. However, they made many comments about the whole process that showed an awareness of Quality Assurance issues. They described problems such as only modelling one building at a time when several may be planned in an area "*changing shape of the city*".

For RADIANCE the realistic appearance and graphic nature of the output is extraordinarily attractive to the design community. It is likely that it and programs like it[15] will be used more and more to provide "pictures" of interiors. However, interpretation of quantitative information from these pictures is problematic. The luminance information obtained from Radiance and other rendering programs is difficult to verify. In daylight applications, the accuracy of the reflectivities of the materials and the geometry of the light reflecting surfaces is not the only limit on the accuracy of the picture. It is also dependent on the accuracy of external data like the sky hemisphere luminance model. Rendering programs currently provide no tools to permit users to calibrate their simulation and ensure its validity - to maintain Quality Assurance.

A further complication arises in the presentation of the results of all simulations. Defining a standard situation is very difficult. For daylight simulation, definition of a "standard sky" is difficult. The question of defining conditions representative of different times of day and times of year is fraught. Even if we could produce a `movie' of each hour of each day for a real year for a daylight simulation, analysing it would be impossible.

Any specification of a simulation tool for architects must include an allowance for Quality Assurance. This has two aspects: ensuring that no one makes simple mistakes in coding data; and more significantly, ensuring that the computer model represents the ideas and reality that they are designing and constructing.

OVERVIEW OF INTERVIEWS

The research interviews showed that in most situations architects do not use environmental design tools. Even when architects seek the predictions of those tools, they apply them at a stage in design when practical improvement in building performance is nearly impossible. There is a mismatch between building performance design tool input/output (i/o) and architects' expectations of their role in that i/o. An associated problem is that the current generation of environmental design tools cannot provide design information in the form that architects would like at these early stages in design. This problem, however, is being addressed by developers of a new generation of computer programs for environmental design analysis[16,17]. Analysts in the surveys reported high expectations of this new generation of program. Program developers are tapping into their wish lists.

However, it still seems questionable whether these programs will ever produce the answers to questions that *architects* want answered at the earliest phases of the design process. This is because most environmental analysis programs are still easier to use in the latter phases of design - the so-called "developed design" phase(s). They help designers to refine their basic ideas, not to formulate the basic ideas.

To build a design tool that helps the designer to formulate a building design that performs well, we must construct tools that function when the building description is vague. The architects interviewed saw themselves as strongly interested in the qualities of the environment that design tools can help characterise.

A few design tools have been developed to provide beneficial "output" from architects' "inputs". Such tools are not new phenomena: Waldram diagrams[18], and R value calculations[19] are well-known examples. They all require designer numeracy. At times education programmes address this need for numeracy by aiming to show how to use design tools "correctly". More often, the education programmes train (pass on skills of use) rather than educate (pass on understanding of principles).

Architecture firms' reluctance to get involved in design performance prediction apparently has its origin in the partners' belief that numeracy skills are not part of their core business. Also, they have no confidence in the skills of junior staff who may have received training in these evaluation techniques. This is because they have no way that they personally can determine the quality of the work done by this junior.

What these same partners find attractive about light rendering software is that, at least superficially, they can use their traditional "architectural" skills to assess the quality of the "output" pictures.

CONCLUSION

Not only for senior partners in architecture firms but for all users of performance prediction software, the greatest single need right now is for Quality Assurance mechanisms. The interviews described above have shown that all users require some means of ensuring that the model they have created with a design tool represents the real building. If performance prediction programs contained good Quality Assurance tools, then architects' interest in the environmental quality of their buildings would naturally drive the use of this software. The problem for these architects is that there is no independent measure, no benchmark to legitimise the output.

We can measure something as simple as an R-value calculation against a benchmark in a code or standard. Yet even for this R-value calculation guaranteeing quality is difficult. No systems exist for independently verifying the calculation, apart from repeating it carefully and comparing the results of the two calculations. The issue is not the precision of each number but the accuracy of the relationship between the numbers and the reality they represent. One cannot easily scan the output and see in it that something is inconsistent or illogical. Improvement of Quality Assurance (QA) procedures for environmental performance prediction will make performance prediction design tools more accessible not only to the professions who currently use them, but also to those architects who currently avoid them.

Design simulation requires building designers to develop a mental model of the relationship between the real world and the information they are feeding into and getting back from the simulation. The quality of this mental model determines the quality of the information that they can obtain from the simulation. If designers do not understand the simulation process, they cannot easily use the simulation results to inform their design. The conscientious but uninformed user will have a series of numbers and a set of concerns about their meaning and reliability. There is an associated danger that the casual but uninformed user will have a series of numbers they trust unreservedly.

It is not easy to imagine a QA process for simple formulae-based simulations such as an R-value calculation or a Reverberation Time (RT) calculation. Quality Assurance of these calculations inevitably degenerates into a process of checking and rechecking the numbers entered into the formulae against their "book" values. Checking what those numbers represent is not so easy.

QA for computer-based calculations still requires that this type of foundation work be done -validation and calibration are always necessary. The formulae and the data values entered must be checked against book values. Validation is normally the role of the writers of the program, and only needs to be done once, when the tool is first compiled. Calibration is required when the software is first set up in an architect's office.

QA in simulation software should allow the user to check the relationship between the performance predictions and the actual building design. A reduction in architects' reluctance to take responsibility for the predictions of simulation software is likely if simulation software produces reports in language used by them.

The principal problem is how to establish a QA system for calibrating the output of a simulation program to ensure its predictions represent reality. What we need is a test for the output from a simulation program like the *Turing Test*[20] for the "existence" of computer-based intelligence. We would apply this test as a theoretical analysis of the output of any environmental simulation program.

Alan M. Turing introduced the Turing Test as 'the imitation game' in his 1950 article where he considered the question "Can machines think?" In the 'modern' version of the test we connect an interrogator to one person and one machine via a terminal. The interrogator's task is to find out which of the two candidates is the machine, and which is human, only by asking them questions. If the interrogator cannot decide within a certain time then we deem that the computer is 'intelligent'.

In the proposed QA test the user must ask whether the data in front of them representing a building's performance is from a real building. Like the Turing Test, this test requires a minimum of three 'players'. Player A asks questions; players B and C answer from their data. B has data on a real building. C also has simulation results. If the interrogator A cannot distinguish simulation from real then the simulation quality is assured.

The problem with this QA idea is the same problem as affects the Turing 'Test': how to operationalise it. Very few offices can afford to have three people working on QA. Also, if real building performance data is available, why simulate? In practice, it is likely the test will involve only one person with the computer taking the other two roles. We suggest that where possible the interrogator should not be the simulationist because it is too easy for self-assessment to become self-congratulation.

In computer-based simulation a post processing program or utility would play the part both of the player with simulation data and of the player who has a 'real' building. Real building data is likely to be a combination of case studies constructed from monitoring programmes in real buildings and from structured parametric runs of the simulation program itself. We need an independent database of this information. It ought to be Internet accessible and able to be added to by submissions from all users. A model could be the CDDB[21]. A unique number system like the Internet URL could be assigned to each building, so that people could upload and download 'cases' from their own set of real buildings. The database need not be too specific. Its primary content would be pointers to the locations for the full data on each case building.

The basic question posed by this QA test is a truism: **The changes in the predictions of a simulation program following changes in building design should always be of the same scale and nature as those perturbations in performance observed in reality.** This obvious "truth" is one that most simulationists would agree is necessary. For example, what use is there for a formula for calculating Reverberation Time (RT) of an auditorium if it only applies to the size of auditorium for which it was derived?

An aspect of simulation that the non analyst often finds puzzling is the external environment. What analysts debate is how to characterise the 'typical' external environment. Is it an average day/week/year? What might the risk to the building owner or operator be if the normally expected variations around the average occur from year to year? Stochastically valid risk analysis is essential in all Quality Assurance procedures related to building performance simulation.

An often-overlooked aspect of the external environment is the operational environment. The designer needs to know just how vulnerable the simulated performance will be to variations in the way we occupy or operate the building. If we no longer operate the building as we assumed it would be, what might the performance consequences be?

The increased complexity of modern computer-based building performance simulation tools has not rid the design profession of its traditional problem with these tools: that they evaluate completed designs. Guidance about how to move forward in improving a design comes only from the informed user looking backwards at how the existing design performs. Post Occupancy Evaluation (POE) contributions to the Internet database could provide a rich set of performance data. We could use it to generate initial design ideas based on successful precedents.

This research defines a development path for the next generation of design tools. It assumes this next generation of design tool will be more detailed computer programs. It also assumes that simulation programs like DOE2 and RADIANCE that a few expert "simulationists" currently use will increasingly be a part of the building designer's repertoire.

REFERENCES

1. Donn, Michael "Simulating Imagined Realities" PhD thesis in preparation, February 1999.
2. Winkelmann, F.C., Birdsall, B.E., Buhl, W.F., Ellington, K.L., Erdem, A.E., Hirsch, J.J., and Gates, S. "DOE-2 Supplement Version 2.1E" (1993)
3. Ward, G.J. 1990 "RADIANCE: A Tool for Computing Luminance and Synthetic Images", Lighting Systems Research Group, Lawrence Berkeley Laboratories, California, USA.
4. Johnson, Paul-Alan "The Theory of Architecture, concepts themes and practices" van Nostrand Reinhold, New York, 1994.
5. Jackson, Anthony. "Reconstructing Architecture for the Twenty-First Century" Univ of Toronto Press, 1995. "It is a telling commentary on the current situation that architects must now be convinced that it is no mean achievement to design buildings that function well, and that allow people to carry on their social life in a practical way".
6. Schön, Donald. "The Reflective Practitioner". Basic Books, Inc., New York, 1983.
7. Gordon, Harry, Min Kantrowitz and Justin Estoque. "Commercial Building Design. Integrating Climate, Comfort and Cost" van Nostrand Reinhold, Co., Inc. New York, 1987.
8. Balcomb, D **Energy 10** in Proc 1997 IBPSA Conf, Prague
9. Gordon, Harry, Justin Estoque and Min Kantrowitz op. cit.
10. Breuer, David, Peter Sterios and Joel Swisher "Design for the Sun - residential design guidelines for New Zealand" Ministry of Energy, 1985
11. Gordon, Harry, Justin Estoque and Min Kantrowitz op. cit.
12. **W**indows, **I**cons, **M**ice, **P**ointing device**S**
13. Breuer et al 1985, op. cit.
14. Marguerite Holloway, "Pattie"; Wired magazine, Dec 1997 (Interview with Patti Maes)
15. Lightscape; Accurender and Microstation incorporate reality based models which permit the user to include real specifications of light sources and fittings in their CAD models. In addition programs like Design Workshop Pro have Radiance Export modules and the vendors even provided an Internet based Radiance rendering service (Artifice Inc. http://www.artifice.com.)
16. Balcomb, Doug. op. cit.
17. Crawley, D.B. et. al. "What Next For Building Energy Simulation—a Glimpse of The Future" Proc. IBPSA 97, Prague, 1997.
18. Waldram, P. J "Universal diagrams" Journal of the RIBA 1933, v. 40, p. 51cited in Solar Control and Shading Devices, V. and A. Olgyay, princeton Univ. Press, Princeton, 1957.
19. ASHRAE "Handbook of Fundamentals" Section 22, ASHRAE, Georgia, 1989.
20. http://www.sscf.ucsb.edu/~sung/comm115/writing-define-computing/Computing-machinery.html *) Computing Machinery and Intelligence (Mind, Vol. 59, No. 236, pp. 433-460)* © Pinar Saygin http://www.cs.bilkent.edu.tr/~psaygin and Varol Akman http://www.cs.bilkent.edu.tr/~akman Created on April 6, 1998 and Last updated: July 7, 1998: http://www.cs.bilkent.edu.tr/~psaygin/ttest.html
21. *CDDB is a net of computers around the world which manage a database with information like artist, title and track title about all available CDs.* From CDCOPY Program Helpfile © Markus Barth (mbarth2193@aol.com). Further information can be found about CDDB at http://www.cddb.com or http://cddb.manawatu.net.nz