VISUAL DATA EXPLORATION IN SUSTAINABLE BUILDING DESIGN

Ralph Evins[1,2], Daniel Knott[1], Philip Pointer[1], Stuart Burgess[2]
[1] Buro Happold, 17 Newman Street, London, W1T 1PD, UK
[2] University of Bristol, Tyndall Avenue, Bristol, BS8 1TH.

ABSTRACT

Visual methods of exploring and communicating data are particularly important to the field of building simulation, where there is always a lot of data but it is not necessarily well-understood or well-presented. Tools have been developed to help address this issue. Three case studies are presented, covering tools that focus on interactive data exploration, animation of data, and multi-objective optimisation data. Each gives feedback on their use on live projects obtained from open-ended questionnaires and informal interviews. It is concluded that such tools can significantly enhance the exploration, understanding and communication of simulation data for sustainable building design.

INTRODUCTION

This paper

In this work, three case studies applying visual data exploration processes to sustainable building design are presented. The three areas covered are interactive, animation and optimisation tools. First, the tool development process and environment are discussed, and the means by which the tools were evaluated. This is followed by the three case studies: for each study, a general description is given, examples are provided, and the feedback results are presented. Finally the conclusion outlines common themes and opportunities for further research and development.

Previous work on visual data exploration

Methods of interactive data visualisation (see Case Study 1) have been investigated by Buja et al (1991; 1996) and by Keim (2002) in relation to data-mining. Temporal data visualisation (see Case Study 2) in both dynamic and static forms has been researched by Müller & Schumann (2003). Multi-objective optimisation data (see Case Study 3) was the focus of an interactive multidimensional exploration tool by Packham et al (2005) that included clustering based on design variables. Other works of note include Wong & Bergeron's history of multi-dimensional, multi-variate visualisation (1997); the summary of visualisation and data mining by Fayyad et al (2002); the detailed guidance on presentation of Harris (2000); the wide ranging Periodic Table of Visualisation Methods (Lengler & Eppler 2007a; Lengler & Eppler 2007b); and the stylistic guidance aimed at achieving both elegance and clarity in many books of Edward Tufte, e.g. (Tufte 2001).

Visual data exploration in building simulation

Modern building simulation produces vast quantities of data. This has led to the need for better methods for understanding the meaning hidden in complex datasets, and communicating this meaning effectively. This contributes to better design space understanding, allowing informed decision-making. It is the job of building simulation practitioners to assist in deriving the most benefit and understanding from simulation data.

There has long been an interest in how to use graphical means to solve problems in building design. This began through architecture, e.g. (Laseau 1975). The use of virtual reality has attracted great attention (Taylor et al. 2007). This is often concerned only with 3D imagery, however may be extended to data display as an overlay (Bouchlaghem et al. 2005). Recent innovations include investigated visualisations of optimisation solutions for architectural design purposes (Caldas 2008), including trade-off front, daylighting and space layout issues, and the "Phi-array" of Mourshed et al (2011), used to explore the design space of a daylighting problem as explored by a genetic algorithm.

TOOL DEVELOPMENT

Process

The tools used here were developed cumulatively over a series of projects. The tool development process was a combination of capability-driven and needs-driven. The ease with which new tools and features could be developed on a timescale suitable for deployment on live projects encouraged new ideas to be trialled as they occurred. Similarly, as new design problems were encountered, the tools were tailored to suit the new challenges that arose.

Development environment

All the tools used here were created using Form Controls and charts in Microsoft Excel worksheets, sometimes with additional VBA scripting. The advantages of this platform include:

- the simplicity with which controls can be configured and arranged
- the existing functionality of Excel (calculations, charting, formatting)
- the ease of automation and advanced control using VBA (even for novice programmers)
- the wide availability of the Excel package (including the availability of VBA codes online)
- the established use of Excel for many data processing tasks and post-processing calculations, allowing integration with the current workflow.

EVALUATION

The tools have been evaluated qualitatively based on user feedback, the aim being to gather a cross-section of a small number of opinions (since development is still early-stage and deployment is not extensive), rather than to conduct a detailed usability assessment (as might be appropriate in later stages). Feedback has been obtained from engineers who applied the tools on live projects and used them to present data to others. All those surveyed had similar relevant experience (use of a tool on a project) and profiles (all were mid- or senior-level engineers). They were selected as the survey group as they were able to:

- comment on capturing technical information in the tools, as they had a high degree of technical knowledge;
- give their perceptions of how well the recipients understood the information, based on the questions and issues raised during use;
- compare the use of the tools with standard design practice, having usually been involved in similar situations in the past;
- suggest further improvements based on their experiences.

Feedback was collected in two ways: via an open-ended questionnaire, and through informal discussions. The questionnaire was sent to all individuals who have used the tools discussed here on live projects. Rather than asking specific questions, respondents were asked to comment on broad topics. Open-ended questions have been noted to avoid bias, and to encourage spontaneous proffering of additional information (Schuman & Presser 1979). Closed-questions, if used to increase survey efficiency, must be based on pilot open responses; this is beyond the scope of this work (though future evaluation may frame closed questions on the basis of the open responses provided here). Suggestions of possible questions to answer were provided for each topic as an example, but it was not compulsory to phrase answers in this way. The format of the questionnaire is given in *Table* 1. The questionnaire also listed details of the project in question, as a reminder to respondents (shown in italics).

Table 1

Format of the questionnaire

Project	*Name of project*
Tool	*Name and type of tool*
User	*All users on this project*
Image	*Screenshot of tool*
Application stories	E.g. When a tool was used on a project, how was it included in the design process? How do you feel the design team responded?
Examples of benefits	E.g. How did this provide an advantage over previous design practise?
Examples of problems	E.g. When difficulties have arisen using a tool, what do you think caused this? How could we improve things to prevent this?
Future work	E.g. What other areas of design practise could benefit from these tools? What other means of data communication could be included?
Other comments	

Response to the questionnaire was a disappointing 6 out of 13 (see Table 2) – busy engineers failed to find the time to provide feedback. Informal interviews with those individuals who did not respond to the questionnaire obtained a better response: 5 (out of 7) interviews were conducted. Such interviews were guided by the issues covered in the questionnaire, to ensure that feedback was comparable between the two methods. The topics (Table 1) formed the basis of the interview structure, and the example questions were used to provide guidance and feedback if necessary. Further interview techniques were taken from (Cannell et al. 1981).

Table 2

Response statistics

Study	**1**	**2**	**3**	**All**	**Ratio**	**%**
Projects	8	2	3	13		
Questionnaires	2	2	2	6	6/13	46
Interviews	2	1	2	5	5/7	71
Total	4	3	4	11	11/13	85

For each case study, the evaluation discussion covers all projects for which feedback was received, with two projects presented as examples. Other feedback not directly related to the tools shown has also been included where relevant.

CASE STUDY 1: INTERACTIVE

Description

The interactive tools presented here provide near-instant feedback in response to changes made by the user. This can be achieved by performing live calculations, by presenting pre-calculated results, or a combination of both. For relatively simple calculations (e.g. average daylight factors or simple energy consumptions), incredibly rapid evaluations are possible, allowing ‘live’ tools that appear to respond instantly. For longer calculations (e.g. detailed shading or full annual simulations) calculations must be performed in advance. Results can then be displayed selectively based on user actions (what Keim refers to as filtering). The final option is to combine pre-calculation with live calculations, for example by interpolating between discrete pre-calculated results. There is a need for caution in the use of interpolation, as with all approximations. However, it can provide seamless user inputs (i.e. avoiding stepped variables) and reduce the number of simulations needed.

User inputs can be discrete variables, using tick boxes or drop-down menus to select preconfigured options. Alternatively users can input continuous variables using sliders. Either can give a much more intuitive ‘feel’ than inputting numbers as text. The outputs included text, simple charts, and 2D and 3D plots. Each element in isolation is very simple: it is how they are combined in a dynamic way that adds value.

Example – effective g-value

The tool shown in Figure 1 calculates an annual ‘effective g-value’, i.e. the ratio of external solar radiation to transmitted solar radiation, measured over the whole year and accounting for all shades, fritting and glazing properties. The tool is based on interpolation between pre-simulated results for four main variables (horizontal and vertical shading lengths; glazing height and width) and also between the four cardinal orientations. Two other variables (glazing g-value and fritting percentage) were used to factor the results. Simulations were run using the IES software package for all combinations of the extremes of the four main variables plus the four cardinal orientations ($4*2^4 = 64$ simulations). Inputs to the tool were the six variables listed above, plus orientation, all via sliders (i.e. continuous variables). The outputs were two-fold: a simple graph and readout for effective g-value, the only numerical results, and a visualisation of a single zone, including North arrow and shades that resize as the inputs are adjusted.

Example – form changes

Figure 2 shows an interactive form tool. The facade was divided into small sections, and data is mapped to each element based on orientation. The user selected a design variant from a drop-down menu, which determined shape and overhangs. The black line shows the roof edge (i.e. the overhang); the dark blue line shows the glazing line. Simulation outputs (from the IES software package) are displayed for each section of the facade: the red line shows the magnitude of the solar gain; the pale blue line shows the daylight penetration into the space. The graph on the right gives the peak solar gain.

Evaluation

Comments often referred to the tools as being “useful in meetings”, for providing “on the spot recommendations”, and of users being “able to see the impact of design changes”. These all praise the instantaneous aspect of the tools. By dramatically reducing the time taken for iterations of design simulations, they enable a live discussion of available options rather than a proposal-evaluation-decision cycle. This also means that all team members can be present for the process, rather than information being exchanged between only a few individuals.

Another comment was that the design team were “keen to play” with the tools. This implies novelty and interest, both good things in teams jaded by too many dull graphs. It also relates to the idea that ‘playing’ can be a critical part of development and exploration, as discussed in Think, Play, Do: “‘Play’ gives shape to ideas, enabling selection, manipulation and learning about possibilities” (Swan 2006; Dodgson et al. 2005). This is linked the “ability to use variable inputs in a very tactile way” – to enable ‘playing’ via these tools, the input mechanisms must feel natural.

Another response was that tools were useful for comparison purposes. This emphasises the fact that data never exists in isolation – context is everything. This could be in relation to targets, as with Figure 1: it aided “comparison with the LEED baseline building”. Alternatively it could be a comparison between different options for a design. Whereas currently designers might evaluate a small number of discrete options, interactive tools allow the entire design space to be explored very rapidly.

Other comments mentioned the power of “complex analysis delivered in a simple, easy to understand form” and “demonstrating the things that were ‘not very sexy’ ... but had real benefits”. Many aspects of building design are more dramatic and tangible than energy performance. Hence this is an area in which great attention must be paid to how information is communicated, to both ensure clarity and simplicity and to engage the interest of the recipients.

The response that “we know some of this stuff inherently but we still need to demonstrate it” shows that these tools, particularly at early stages, can embody knowledge already held by experts, but in need of communication to everyone else. Conversely, another response mentioned the risks of “giving sophisticated tools to people that don’t fully understand the nuances of them”. This highlights the need to design tools to avoid this problem, and to be clear about all underlying assumptions.

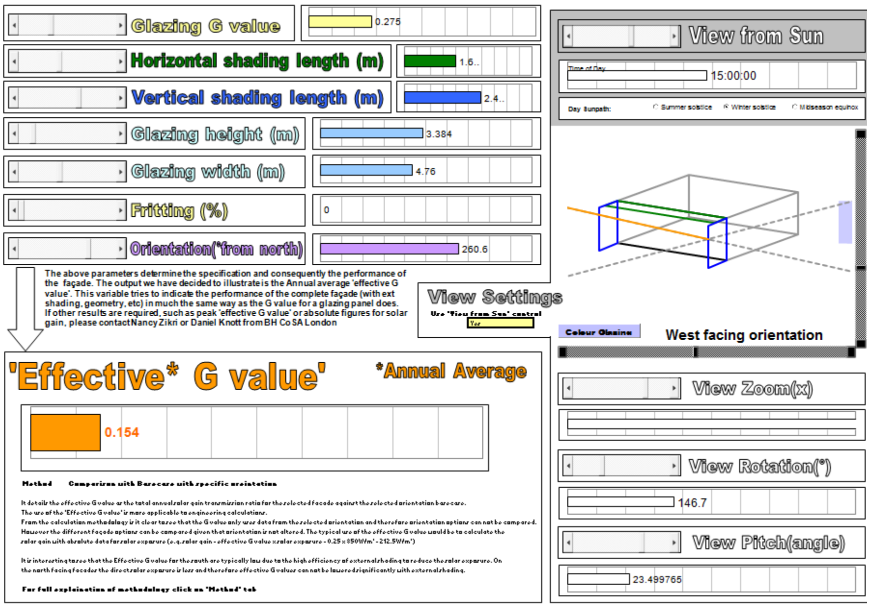

Figure 1
Interactive tool example: effective g-value

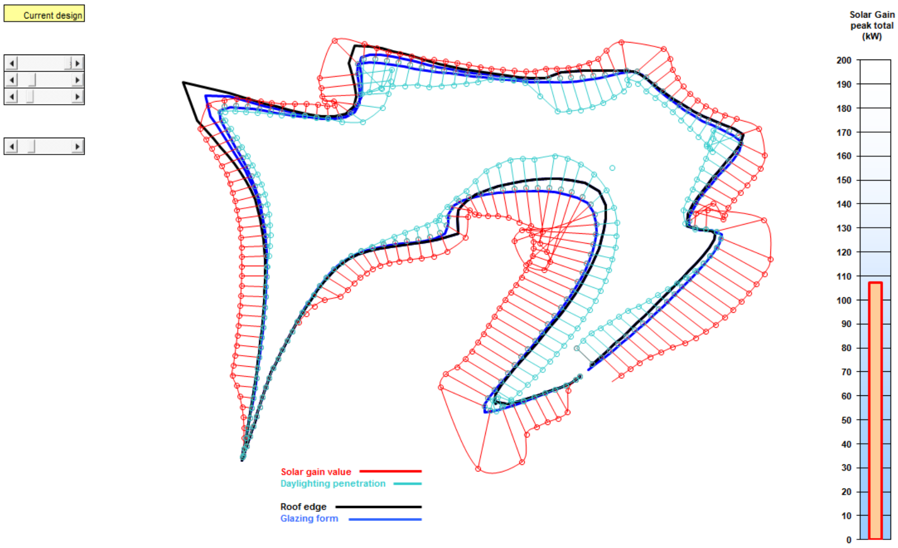

Figure 2
Interactive tool example: form adjustment

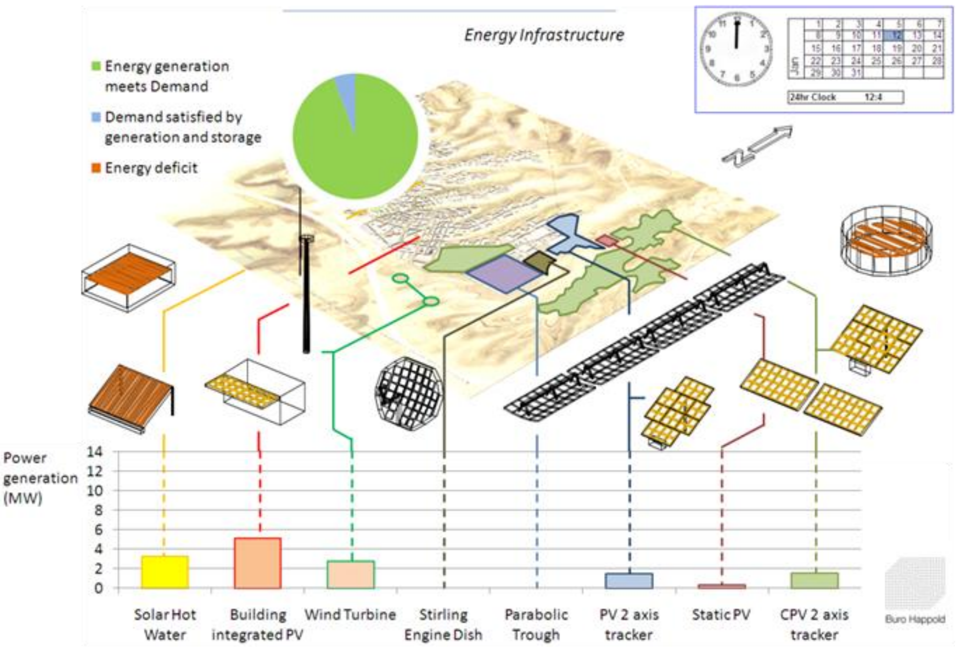

Figure 3
Animation tool example: energy mix

CASE STUDY 2: ANIMATION

Description

For some applications where there is a strong temporal component to the data, animating the visual as a video gives a clear view of the different states, how they develop and interact, and how they relate to targets or criteria. This has been used extensively for inherently time-based processes like displaying shadow patterns and sun paths. However, it is not limited to such direct processes: any temporal data can be viewed as an animation.

Whether this is useful or not depends on the information to be conveyed. In some cases, a timeseries graph is a much better way to display the information - trends, correlations and interdependencies can all be distinguished much more easily if the time history is shown rather than a single instant. However, animations may be useful when the desire is to communicate complicated information as clearly as possible, and time dependencies are less critical to the message. Animation can provide a less cluttered visual by removing the unnecessary historical data, whilst still allowing different times and states to be shown. Alternatively, it is possible to combine an animation with a timeseries plot, including a moving highlight of the current time (e.g. Figure 4).

Animations can be easily integrated with the means of presenting ideas in meetings or to an audience. While fixed graphics are intended for print, animations make sense when incorporated into presentations. A speaker can narrate an animation, highlighting different facets of the information as required. Indeed, modern presentation methods are moving away from slides (a sequence of discrete units) towards the use of animation and video (a continuous, fluid medium), either by including clips or by playing a single video. With this in mind, animated outputs of simulation data can incorporate many advanced techniques borrowed from cinematography, for example fades, pans and zooms.

Example – energy mix

The snapshot in Figure 3 is from an animation to show the mix of different sources in a renewable energy project. The data was calculated from a weather file, applying quick calculations to derive the energy available from solar, wind etc at different times of the day and the year. The main graph shows the power input from each source. These are illustrated with a graphic and linked to zones on a site plan. The calculations were conducted based on the chosen equipment capacities. The pie chart (upper left) shows energy supplied, energy from storage and any deficit in supply. The passage of time is also shown graphically via an analogue clock and calendar.

Example – temperature gradients

Figure 4 shows a snapshot of an animation illustrating temperature gradients in a building. Some areas (the outdoor atmosphere and the earth tube labyrinth) have uniform temperatures, while other areas (the atrium and the individual zones) have temperature gradients across them. Temperatures were read from a data file for each timestep and converted to the colour scale shown on the right. For areas with temperature gradients, a temperature was taken for either end, and a linear colour gradient applied. The zone temperature is also shown on a timeseries graph at the bottom (along with thermal comfort bands), and a red circle marks the current point. This clearly indicates when overheating occurred; when it did, a graphic appeared to indicate zone cooling, and the temperature trace on the lower graph followed the cooling set point.

Evaluation

Benefits reported regarding the use of animated tools on live projects included the ability to “impress [the] client team” and to effectively “articulate the ... story”. This sums up the power of such tools: data presentation can become impressive, and can be communicated articulately.

Problems included “quite a lot on screen at the same time”. Animated data presentation can suffer from the same issues as static presentation: graphics can easily become cluttered with what Tufte defined as “chartjunk”. In addition, because everything in an animation may be constantly changing, it becomes necessary to give the viewer more time to appreciate the significance of each item. One suggestion for how to remedy this was to “step through” the information.

Suggestions for future work built upon what has been developed to date: publicising and disseminating the established tools. It was also suggested that integrated building and systems modelling would benefit greatly from this approach - this is another area in which. interrelated temporal data is prevalent.

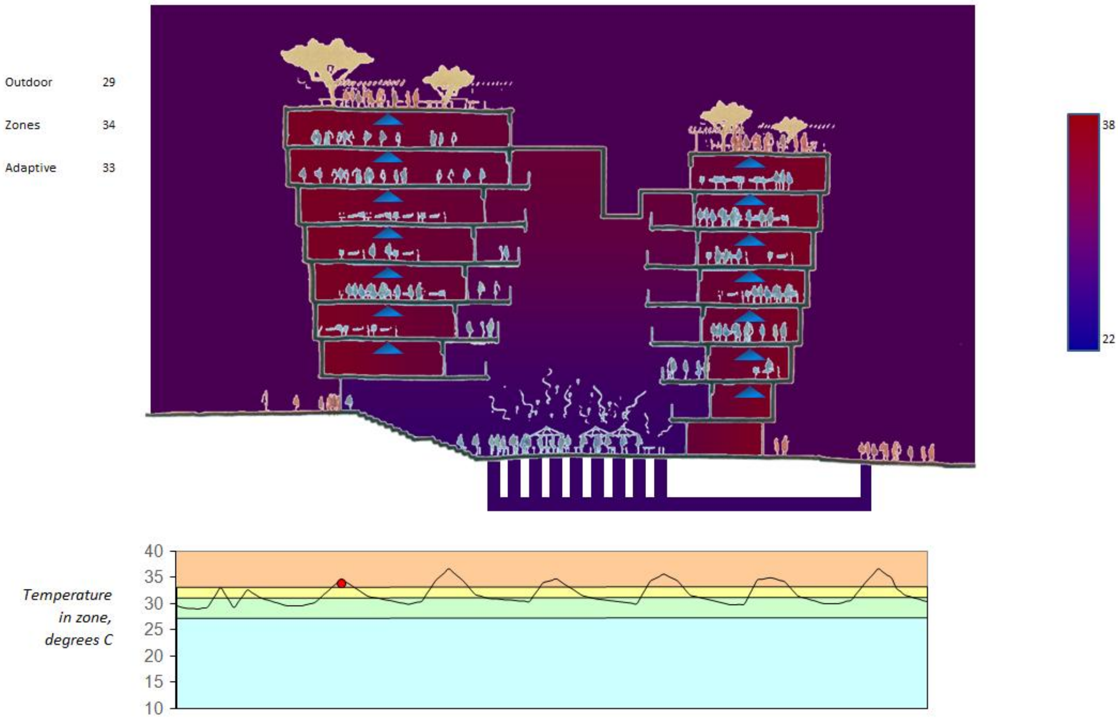

Figure 4
Animation tool example: temperature gradients

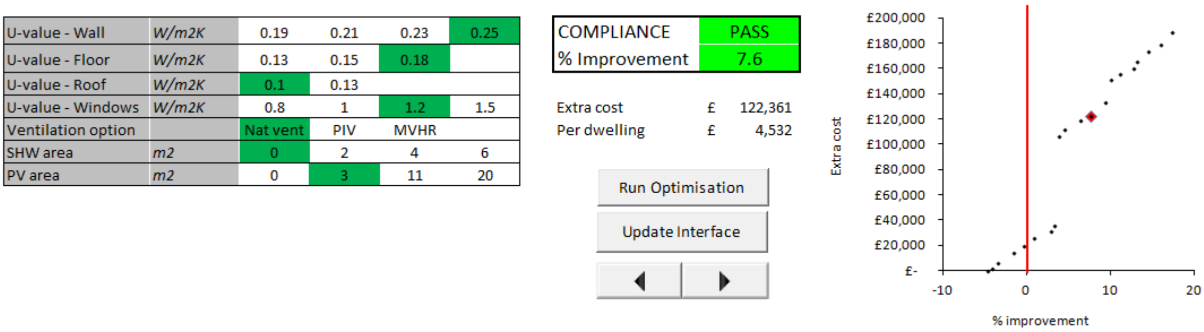

Figure 5
Optimisation tool example: solution selection

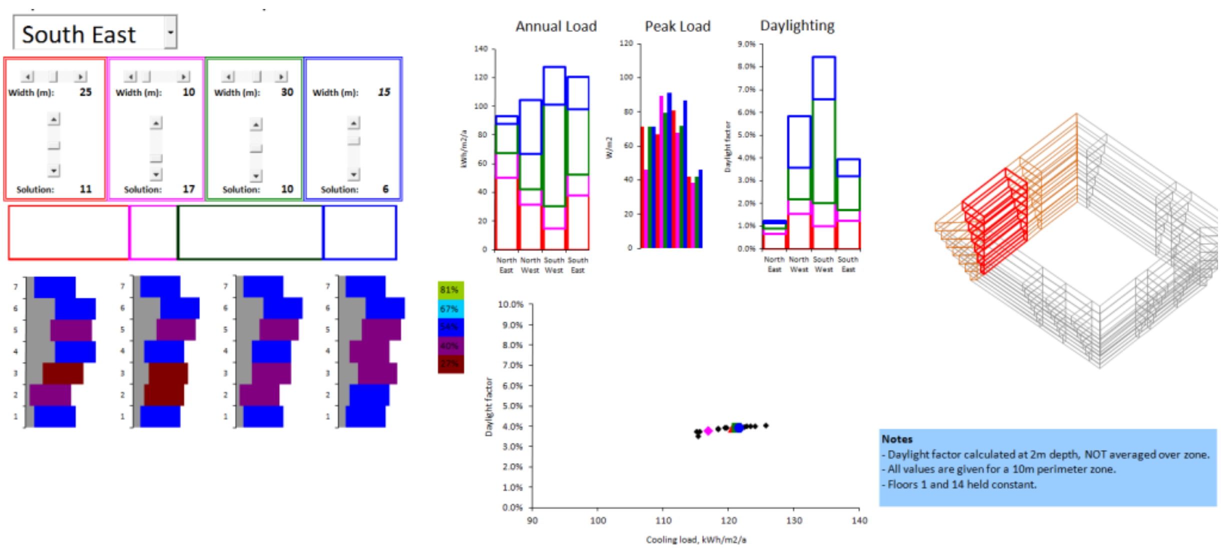

Figure 6
Optimisation tool example: facade profiles

CASE STUDY 3: OPTIMISATION DATA

Description

Multi-objective optimisation (Coello et al. 2007) uses computational optimisation methods to find a set of solutions that form a trade-off front. Solutions on this front are referred to as non-dominated: there are no solutions that perform better in all objectives. All solutions on the front are optimal for some balance between the objectives. The algorithm aims to find solutions as near to the true front as possible, and distributed evenly along it.

Multi-objective optimisation presents significant challenges in visual data exploration. There are at least two objectives that must be plotted against each other to see the shape of the trade-off front. If there are more objectives, visualising the trade-off hypersurface becomes difficult. For each solution (a single point in objective space), there is a corresponding set of values of the design variables. The design space dimensionality (the number of variables of the problem) is often high – rarely less than five variables, maybe twenty or more.

There are static means of presenting these complex interactions. For example, 'innovisation' (Deb & Srinivasan 2006) uses many simple plots to discern relationships amongst optimal solutions. Plotting variables against individual objectives for optimal solutions examines a slice along the trade-off front. However, requires a large number of such plots. An alternative is to combine this information into a single graphic, as in (Evins et al. 2012) which used colours for each variable, allowing the reader to distinguish patterns such as blocks and repetitions. It also omitted unused information and arranged the data in bands so that the graphic could be read vertically, similar to a tree diagram. The result is information-rich but not particularly intuitive. The tools presented here address these issues using interactive means, as presented in the first case study.

Example – solution selection

Figure 5 shows a simple tool used to select an individual solution from an optimised set. The optimisation problem concerns the trade-off between carbon emissions (measured as the percentage improvement over a baseline) and cost increase. The graph on the right shows all optimal solutions, with the red marker highlighting the one selected; arrows are used to change the selection. The vertical line indicates the target performance. Details of the chosen solution are listed on the left: values of each design variable, plus objective values and a pass/fail indicator (which also changes colour).

Example – facade profiles

Figure 6 shows a more advanced tool in which each facade section can be selected from an optimised set. Each solution gives the offsets (set back or protruding) and glazing area (indicated by colour), and corresponds to a certain daylighting and solar gain performance trade-off. A solution can be selected for each of four facade sections (red, pink, green, blue); the section widths can also be altered. This process is performed separately for four facades. Displayed for the current facade are each profile (lower left), section widths (middle left), and solution space with the selections highlighted (lower centre). Also shown are annual solar, peak solar and daylight results for all facades, broken down by section (upper centre), and a 3D wireframe of the whole building (right), highlighting current facade (orange) and section (red, based on most recent adjustment).

Evaluation

Feedback regarding optimisation tools in general highlighted significant barriers to understanding the underlying principles, which hinders effective use of the results. A balance is needed between overloading a non-technical audience with details of the process, and providing too little information so the benefit is not recognised. Whilst not directly related to these tools, this affected their use in practice.

The simplicity of the solution selection tool was praised; it was preferred to static graphs giving the same information. Manually proceeding along the front gave the opportunity to observe associated variable changes. A suggested improvement was for graphical presentation of variable data.

Feedback was generally not positive for facade profile optimisation tool. It tried to do too much at once; the tool attempted to capture a huge volume of information in one place, and failed to maintain clarity. Multi-objective optimisation inherently relates to various disciplines, in this case architecture, facade design and HVAC system engineering. Input from individuals with different interests led to many different features being added. Overall, this led to great technical breadth but poor usability in practice.

CONCLUSIONS

The three case studies presented here summarise the application of visual methods to complex data sets in new ways. They have not been developed in isolation or to fixed requirements, but have evolved gradually in response to the challenges of live projects. User feedback and evaluation has shown that in most cases they have been useful, by better exploring the meaning of data, communicating that meaning, or provoking interest in the message they deliver.

Few suggestions for improvement went beyond refining current tools – this highlights the difficulty of conceiving novel means of exploring and communicating data. Possibilities for new tools include combining existing ideas (e.g. hybrid interactive and animated), exploring other types of visualisations, and incorporating technical enhancements (clickable interactive 3D graphics, more advanced animations). However, it is important to retain ease of development and use to ensure application on live projects.

ACKNOWLEDGEMENTS

Funding for this research has been provided by Buro Happold Ltd and EPSRC. Ralph Evins is a research engineer with Buro Happold and with the Industrial Doctorate Centre in Systems, Universities of Bristol and Bath, UK.

REFERENCES

Bouchlaghem, D. et al., 2005. Visualisation in architecture, engineering and construction (AEC). *Automation in Construction*, 14(3), pp.287–295.

Buja, A. et al., 1991. Interactive data visualization using focusing and linking. In *Visualization, 1991. Visualization '91, Proceedings., IEEE Conference on*. pp. 156 –163, 419.

Buja, Andreas, Cook, D. & Swayne, D.F., 1996. Interactive High-Dimensional Data Visualization. *Journal of Computational and Graphical Statistics*, 5(1), pp.78–99.

Caldas, L., 2008. Generation of energy-efficient architecture solutions applying GENE_ARCH: An evolution-based generative design system. *Advanced Engineering Informatics*, 22(1), pp.59–70.

Cannell, C.F., Miller, P.V. & Oksenberg, L., 1981. Research on Interviewing Techniques. *Sociological Methodology*, 12, pp.389–437.

Coello, C.A.C., Lamont, G.B. & Veldhuizen, D.A.V., 2007. *Evolutionary algorithms for solving multi-objective problems*,

Deb, K. & Srinivasan, A., 2006. Innovization: innovating design principles through optimization. In *Proc: Genetic and evolutionary computation conference*.

Dodgson, M., Gann, D. & Salter, A., 2005. *Think, Play, Do: Technology, Innovation, and Organization*, OUP Oxford.

Evins, R. et al., 2012. A case study exploring regulated energy use in domestic buildings using design-of-experiments and multi-objective optimisation. *Building and Environment*, 54(0), pp.126–136.

Fayyad, U.M., Wierse, A. & Grinstein, G.G., 2002. *Information visualization in data mining and knowledge discovery*, Morgan Kaufmann.

Harris, R.L., 2000. *Information Graphics: A Comprehensive Illustrated Reference*, OUP USA.

Keim, D.A., 2002. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1), pp.1–8.

Laseau, P., 1975. *Graphic Problem Solving for Architects and Builders*, Reed Business Information.

Lengler, R. & Eppler, M.J., 2007a. Periodic Table of Visualization Methods (www.visual-literacy.org/periodic_table/periodic_table.html).

Lengler, R. & Eppler, M.J., 2007b. Towards a Periodic Table of Visualization Methods for Management. *Proc: Conference on graphics and visualization in engineering*, pp.1–6.

Mourshed, M., Shikder, S. & Price, A.D.F., 2011. Phi-array: A novel method for fitness visualization and decision making in evolutionary design optimization. *Advanced Engineering Informatics*, 25(4).

Müller, W. & Schumann, H., 2003. Visualization for modeling and simulation: visualization methods for time-dependent data - an overview. In *Proc: Winter simulation: driving innovation*. WSC '03. pp. 737–745.

Packham, I.S.J. et al., 2005. Interactive visualisation for decision support and evaluation of robustness—in theory and in practice. *Advanced Engineering Informatics*, 19(4), pp.263–280.

Schuman, H. & Presser, S., 1979. The Open and Closed Question. *American Sociological Review*, 44(5), pp.692–712.

Swan, J., 2006. New ways of innovating. *Building Research & Information*, 34(4), pp.427–432.

Taylor, M. et al., 2007. User Interfaces for Complex Engineering Software: Application of Virtual Reality. *Journal of Computing in Civil Engineering*, 21(1), pp.59–62.

Tufte, E.R., 2001. *The Visual Display of Quantitative Information* 2nd ed., Graphics Press USA.

Wong, P.C. & Bergeron, R.D., 1997. 30 years of multidimensional multivariate visualization. *Scientific Visualization, Overviews, Methodologies, and Techniques*, pp.3–33.