

Multivariate Interactive Visualisation of Data for Multi-Objective Design Space Exploration

Andre CHASZAR^{*b}, Peter von BUELOW^a

^{*}O-Design Research & Consulting
ac@0-design.com

^a University of Michigan

^b Technical University of Delft

Abstract

This paper describes recent developments in an ongoing research effort to develop a method, ParaGen, combining parametric form generation with multi-performance optimization and exploration. Here we focus on how the system is augmented with data visualisation capabilities which enhance the understanding of relationships between design parameters and performance, and thus aid the exploration of design alternatives. These capabilities are applicable to a very wide variety of designs, including spatial structures, and we illustrate their use with an example of a long-span gridshell dome with central oculus, having variable parameters for defining the dome's and framing's basic geometries. We demonstrate how various kinds of data visualisations facilitate overviews and checking of the large amounts of data generated by the system, and we conclude that these visualisations improve the comprehension of important relationships within the data as well as guiding exploration of design alternatives.

Keywords: data visualisation, design space exploration, gridshell, interactive visualisation, multi-criterion, multi-disciplinary, multi-objective, optimisation

1. Introduction

Multi-objective design situations arise frequently in work on shell and spatial structures -- as well as many other types -- due to these structures' various, at times complementary but sometimes conflicting performance requirements. Recursive optimisation routines and other generative design processes aid multi-objective design by producing many variations with given design parameters, making many alternative designs available for comparison. However, the potentially very large number of design alternatives produced, and the even larger number of variations possible, both present formidable challenges to the designers' comprehension. Whereas understanding the details of one or a few design options -- as in conventional design processes -- is tractable, expansion of the options' number to dozens, hundreds or thousands (or even more, given sufficient computational resources) requires that significant support be available, if they are all or even mostly to be examined and understood. In conventional optimisation routines, this problem is circumvented to some degree by focusing only on the optimal solutions or a few of the top performing ones. Yet this is not entirely satisfactory, as: 1) the correctness of the analysis is difficult to check with so few samples; 2) much valuable knowledge about the behaviour of the design model is potentially lost; 3) the fixed objectives needed for such a closed optimisation process are not present in earlier design stages; and 4) designers may wish to interact with the process rather than just accept the final outcomes of a 'black box'.

The ParaGen system (or method) addresses these issues in various ways, primarily by:

- a) storing all design data and supplemental files produced (including their input parameters, performance values, images and other files) for all solutions, and making them available for retrieval;
- b) allowing designers to select design alternatives to be further developed, including directly specifying input parameters if so desired;
- c) making it possible to change the objective (or “fitness”) function(s) during the process of design generation without restarting the process or losing progress; and
- d) providing data visualisation tools to help designers make sense of the results.

In this paper we focus on the data visualisation capabilities, and aside from a brief summary of ParaGen’s main features, we refer readers to other publications for more details of other aspects. In the following sections we describe the features and operation of the data visualisation tools and illustrate their use with an example gridshell structure (Figure 1). We close with a discussion of the system’s current performance in terms of the points mentioned above, indicate directions for future work on further developing the system, and conclude that based on our experiences so far, as well as assisting with trouble-shooting and checking of design models’ consistency, the data visualisations do help with comprehension of relationships between design parameters and performance indicators. (Note that illustrative figures are reproduced here at small size due to limitations of space; please contact the authors for access to a website with higher resolution images.)

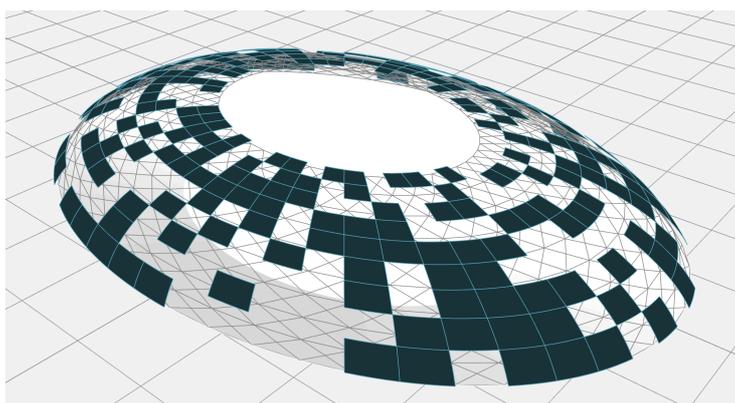


Figure 1: perspective view of example gridshell roof with partially opaque cladding.

2. System

Development of the ParaGen system has been an ongoing effort at the University of Michigan for several years. An overall description of the method can be found in an IASS Journal article from 2012 (von Buelow [22]) as well as several more recent updates at conferences on specific topics such as designer interaction (von Buelow [23]), multi-objective fitness functions using database queries (von Buelow [24]) and data visualisation techniques for design exploration (Chaszar [3], [4], Chaszar *et al.*[5]).

ParaGen is termed a system or method because it actually combines several programs running on different machines with interactive control and visualization interface using a web site. Figure 2 shows the key components of the system. All of the solutions generated and analysed are maintained in a SQL database (5). The database contains both the parameters used to generate the solutions (parametric input values) as well as the performance values which result from one or more digital simulations. In addition any number of files or images can be linked to the solutions held in the database. Ultimately the database holds this solution data in a form which can be accessed in a variety of ways both for selection in breeding more solutions as well as for post-processing analysis and

exploration. The server which holds the database also runs a genetic algorithm (Grierson [7], Miles [13]) to select and breed (1 and 2) new solutions based on parent solutions taken from the database. The new child solutions are downloaded to client PCs (running in a parallel network) in the form of parametric variables. These variables are then processed and the results analysed by appropriate parametric and simulation software (3 and 4). In the example used for this paper, a Grasshopper script was used for the parametric processing and STAAD.Pro provided the structural simulation with a finite element analysis. Finally the performance results of the simulation(s) are uploaded to the database on the server. A web site provides access to the database, and is used to display and explore the solutions in a variety of ways.

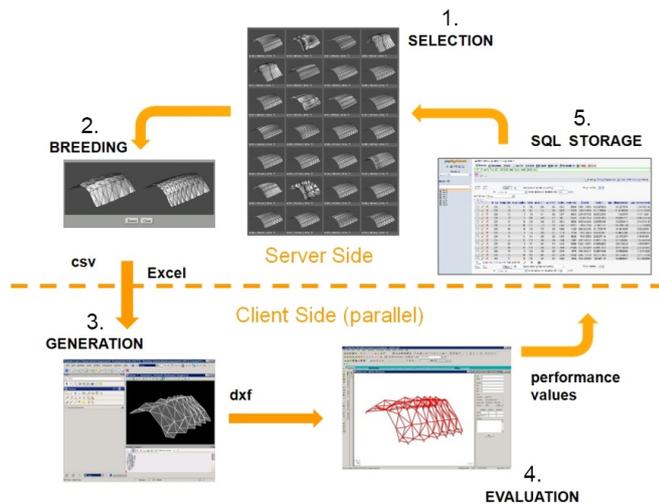


Figure 2: The five steps in the ParaGen cycle. 1. Selection of parents, 2. GA breeding through crossover, 3. Generation of the child solution, 4. Evaluation of the solution, 5. Saving all data in a relational database.

3. Data visualisations

Years of research and practice in the area of data visualisation have resulted in a great variety of visualisation methods (Keim *et al.* [10], Mourshed *et al.* [15]; Unwin *et al.* [19]), a comprehensive review of which is beyond the scope of this article. Instead, we focus on three particular types of visualisations implemented (via D3, Bostock) in the ParaGen system: 1) design images with captioned data; 2) scatterplots (Elmqvist *et al.* [6]; Mayorga and Gleicher [12]); and 3) ‘parallel coordinates’ (or ‘parallel axes’) plots (Alsakran *et al.* [1]; Inselberg, [9]). The aim of each of these is to promote comprehension of the data associated with the various design alternatives generated by providing graphical information supplementary to the alphanumeric tabulation of data which is typically predominant. The graphical information enables comparison of alternatives’ data by visual means more powerful than the viewing of just words and numbers. In the case of design images, the forms of the designs as well as supplemental colouring can quickly convey key similarities and differences. Likewise, plotting of data points and supplemental colouring can summarise similarities and differences for even more alternatives in less space than a ‘gallery’ of design images, although with loss of richness in terms of the number of data dimensions presented in a given scatterplot. Parallel coordinate plots, on the other hand, preserve more (almost all) dimensions of the data. The potentials of such techniques are being exploited in other disciplines (Arentze [2]; Monks *et al.* [14]; Woodruff *et al.* [21]) but apparently only rarely in architectural design and engineering.

4. Example model

We demonstrate the application of these visualisation methods through an example exploring design alternatives for a long-span roof with an elliptical base and central oculus. While the base plan of the

roof is fixed, design variations can occur in: the oculus' size; the oculus' height; the roof's sectional profile; and the subdivision of the roof surface into framing elements constituting a gridshell. (See Figures 1 and 3.) Each of these design characteristics is alterable by changing the values of variable parameters in a parametric-associative model describing the roof's geometry. Each variable parameter is initially defined over a range of minimum and maximum values. While some of these could be continuously variable, for purposes of the example they also have defined step sizes. (This is in any case appropriate for initial design explorations, in contrast with later optimisations where finer gradations of parameter values may be needed.) See Table 1 for parameters, ranges and step sizes.

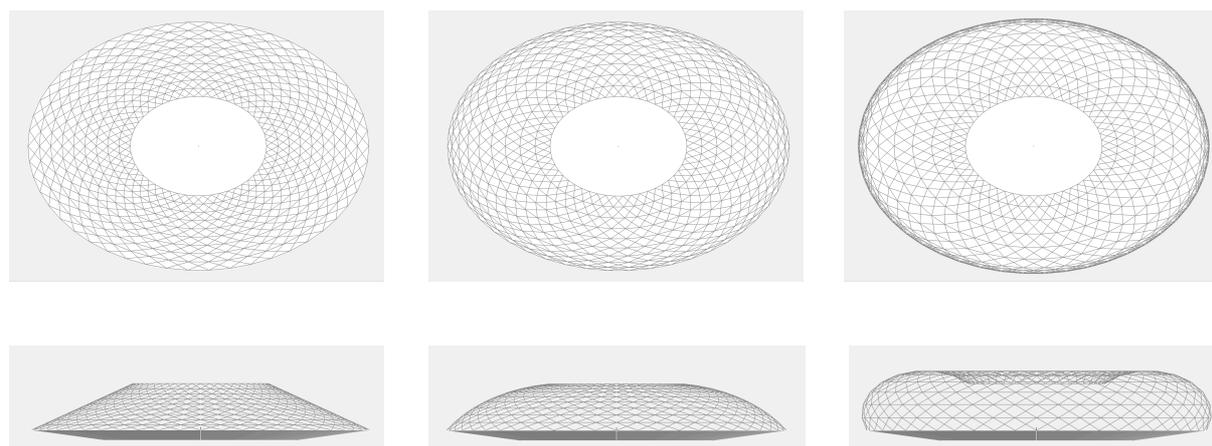


Figure 3: plans (top) and sections (bottom) of roof framing with variation of roof profile parameter from minimum (left) to maximum (right) values.

Table 1: Gridshell model variable parameters for generating alternative roof designs

parameter	range (min - max)	step	remarks
oculus size	5 - 20	0.5	radius (nominal)
oculus height	5 - 20	0.25	(nominal)
profile curve	0 - 15	0.5	offset from line: base to oculus edge
radial divisions	5 - 15	1	concentric rings
circumferential divs.	16 - 50	1	radial spirals

NOTE: values are with fixed roof base nominal size of 55 m (minor radius)

5. Test results

Our testing of the ParaGen system's generative and data visualisation capabilities on the example model presented above comprised general exploratory activities as well as more goal-directed ones. The initial phase of design space exploration consisted of 200 randomly generated designs (fitting within the parameter ranges above) to give an overall sense of the space's characteristics, and also to provide material for further generation via recombination. Scatterplots of various performance criteria and input parameters allow visual confirmation of the ranges covered, as shown in Figure 4 below. Subsequent phases of design generation aimed to minimise or maximise relevant inputs and outputs in order to explore the extremes available within the given parameter ranges. (See Figure 5.) Again scatterplots enable a rapid overview and checking of the values produced. Visible jumps in the plots occur each time the fitness function is altered. This shows the immediate and effective shift in the areas being explored.

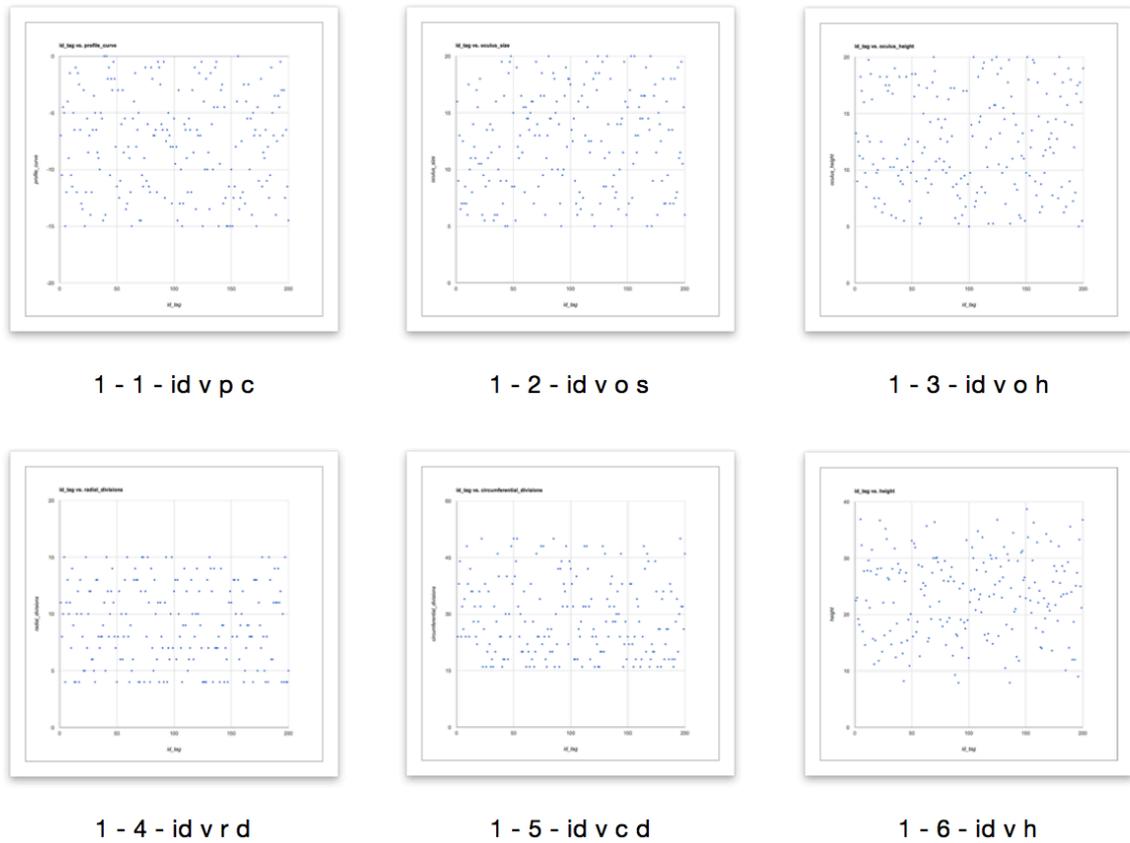


Figure 4: scatterplots showing distributions of parameters for initial population of 200 random designs

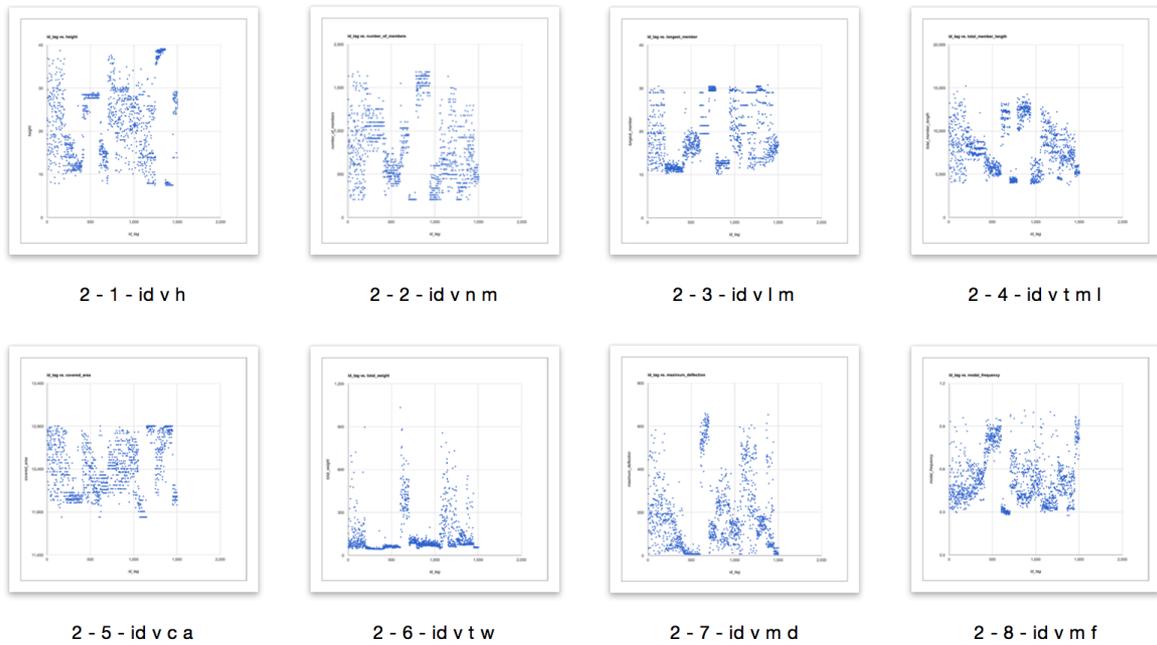


Figure 5: distributions of performance criteria for initial random designs and later single-objective optimisations; criteria shown are (top, left to right): roof height, number of members, longest member, total member length, and (bottom, left to right): covered area, total weight, maximum deflection, modal frequency.

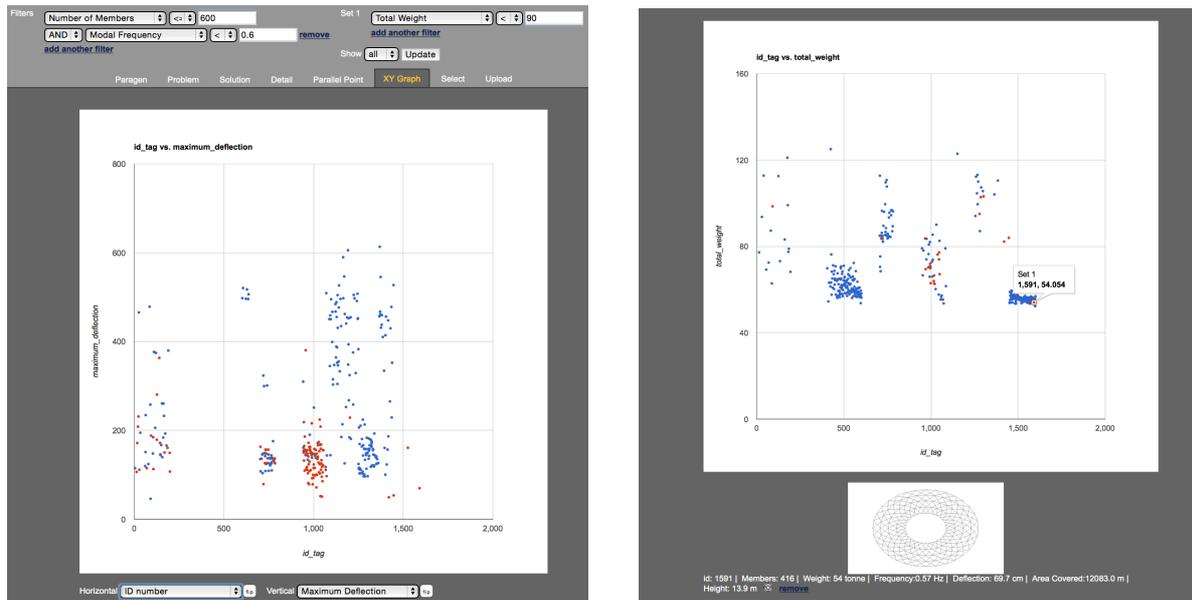


Figure 6: (left) scatterplots filtered and highlighted with user-defined criteria via menus and text fields; (right) clicking the dots shows further design performance data with an image of that solution.

The last phases of design generation aimed to find balanced solutions by using multi-objective fitness functions as follows:

- 1) 600 or fewer structural members; 0.6 Hz maximum modal frequency; 90 t maximum weight; and minimise deflection;
- 2) number of structural members 600 or less; deflection less than 100 cm; and minimise weight.

These two multi-objective search examples are essentially described as SQL queries that define a limited set of solutions which form the breeding population. In this way further child solutions are found by the GA that match these multi-objective criteria. Design solutions meeting these criteria are shown in scatterplots via user-defined filters and highlighting, as shown above (Figure 6). Interactive selection allows on-demand examination of more details concerning particular designs of interest by simply clicking the dots on the plots. Note that some good (non-dominated) intermediate (non-extreme) design solutions are also found within the initial randomly generated set, of course, as well as among others resulting as by-products of the shifts in objective functions.

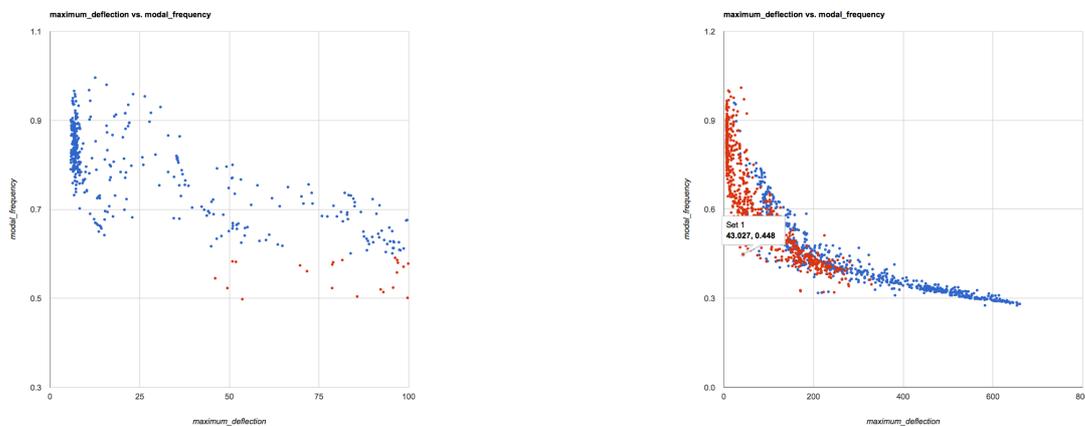


Figure 7: scatterplots showing inverse nonlinear relationship of modal frequency to maximum deflection.

More detailed analysis of the generated designs is facilitated by parallel axis plotting, which simultaneously shows all parameter values for all generated solutions -- or user-selected subsets of those (Fig. 8). This allows examination of relationships among multiple parameters, and highlights the extent to which particular parameters depend on others, or on combinations of others. For the multi-objective criteria stipulated above, parallel plotting gives some sense of the design parameters most influential in affecting the particular balance of performance goals. Automatic scaling of the various axes displays the range of minimum to maximum values for each input and each output parameter. Within these ranges, the extent of concentration or dispersal of values is visible. Interactive highlighting via selection fences facilitates examination of how increases or decreases in particular values relate directly, inversely or otherwise to changes in other parameters. This dynamic selection mode also shows how expansion or shifting of parameter ranges may give access to other designs with even better performance values. For example, allowing an increase in the number of members from 600 to ~700-750 gives access to designs having both low deflections and low modal frequencies, albeit at some cost in weight (see Figure 9).

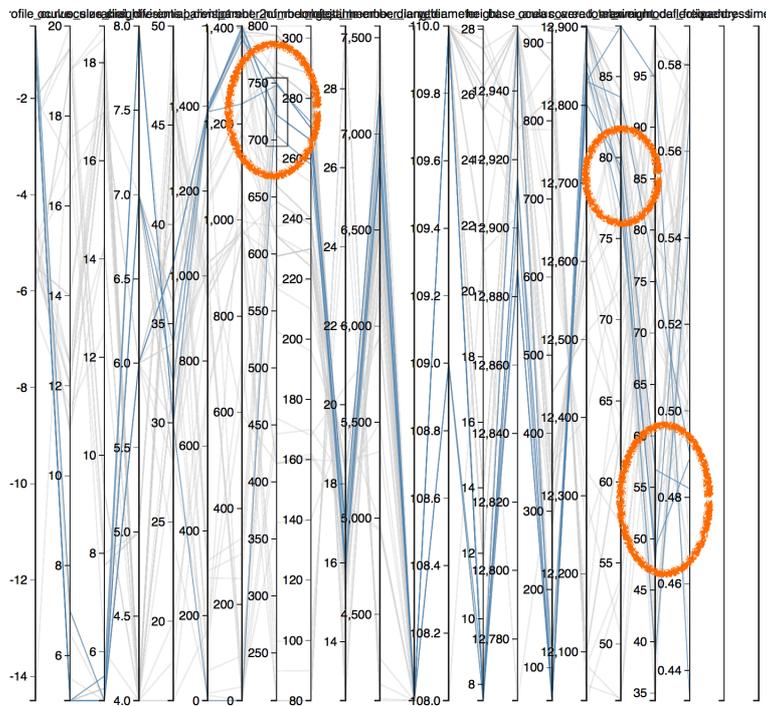


Figure 9: Parallel axis plot showing possible improvements of modal frequency and maximum deflection with increase in range of acceptable number of members, but still within given weight objective.

By shifting as desired between model views, scatterplots and parallel plots, users of the system can thoroughly explore the pool of designs already generated, derive understanding of the relationships governing behaviour of the design model, and guide further generation of designs to explore design space areas of interest, or even transform the design space by changing parameter ranges or changing the parametric model itself.

6. Discussion

As mentioned in the introduction, the methods described above are applicable to a wide range of design problems including the evaluation of shells, gridshells and space frame structures for multi-criterion (multi-objective) performance. Such analyses may also include various non-structural performance considerations (Turrin *et al.* [17], [18]) -- e.g. lighting, ventilation, acoustics, etc. -- which often need to be balanced with structural aspects of the design, resulting in ‘sub-optimal’

solutions that nevertheless can offer better overall performance or at least point to better ones (Sosa & Gero [16]).

While other approaches to such multi-objective design situations are more common in practice -- such as intuition- and judgement-led ones focusing on one design or a few alternatives, case-based reasoning, or 'satisficing' -- due to constraints of design schedules, budgets and skills, the approach we have described facilitates generating and evaluating large numbers of alternative structures (or designs) with variations of parametric model parameter values affecting various scales (e.g. overall form, subdivision, specific design features, etc.) (Hatchuel *et al.* [8]; Woodbury & Burrow [20]) This can enable broader and more in-depth consideration of alternatives even within tight time constraints. Confidence in the results is promoted by interactivity both in supporting users' guidance of the design generation process -- e.g. by specifying changeable objectives (Maher & Poon [11]) and selecting designs of interest for further development -- and in supporting interactive assessment of performance (output) indicators' sensitivity to various design (input) parameters. This in turn enables checking of model behaviour against expectations for validation, and helps comprehension of which parameters to focus upon and which to de-prioritise (e.g. by choosing fixed values for them or limiting them to few variations) as well as where altering their ranges could be appropriate. The result is a design process promoting greater potential for innovation as well as quality of designs within reasonable timeframes.

Further improvement of this approach can be envisioned on some points including: a) a need for systematic exploration of design space as well as stochastic generation, to better build confidence in model behaviour; b) inclusion of statistical analyses to support examination of parameter interrelationships; and c) extension of multivariate interactive visualisation capabilities to enable more effective identification of relevant patterns in the design data, such as automated and interactive methods for dealing with overplotting in areas of high data density.

7. Summary / Conclusion

We have described how the ParaGen system for generative design is augmented with data visualisation capabilities which aim to support designers in understanding the relationships between design parameters and performance indicators, especially for complex, multi-objective design situations. We have illustrated the system's operation and data visualisations with an example model of a gridshell roof having design parameters controlling the overall roof form and its subdivision into individual framing elements. By providing multivariate interactive visualisation tools, the system enables interactive exploration of Pareto-optimal as well as sub-optimal design alternatives, which makes explicit to the designer trends and other relevant information about parameters and performance. We have shown how this helps comprehension of parameters having greater and lesser effects on performance, as well as aiding trouble-shooting and model consistency checking. On this basis we hope to further develop the capabilities of ParaGen and also see interactive data visualisation become a more common complement to optimisation and generative design tools in general.

Acknowledgements

We wish to thank Taubman College at the University of Michigan for the use of the Hydra Lab for Parallel Computing and the support of the ParaGen project, and Michela Turrin for pointing out some references on data visualisation. Kudos also to Mike Bostock for D3.

References

- [1] Alsakran J., Alhindawi N. and Alnemer L., Parallel coordinates metrics for classification visualization. *2016 7th International Conference on Information and Communication Systems (ICICS)*. IEEE, 2016.
- [2] Arentze T.A., Spatial data mining, cluster and pattern recognition, *International Encyclopedia of Human Geography*, 2009, 325–331.

- [3] Chaszar A., Bridging the gap with collaborative design programs, in: *Blurring the Lines*. Wiley-Academy, 2006, 158-167.
- [4] Chaszar A., Finding rhythms in a tempest, in: J. Burry (ed.) *Designing the Dynamic: High-performance Sailing and Real-time Feedback in Design*, 2013, Melbourne Books, 118-123.
- [5] Chaszar A., von Buelow P. and Turrin M., Multivariate interactive visualization of data in generative design, in: *Proc. SimAUD 2016*.
- [6] Elmqvist N., Dragicevic P. and Fekete J-D., Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE transactions on Visualization and Computer Graphics* 14.6, 2008: 1539-1148.
- [7] Grierson D.E., Conceptual design using emergent computing techniques. *Emergent Computing Methods in Design; Applications of genetic algorithms and neural networks*, NATO ASI Series F, Computer and Systems Sciences, 149, Springer, 1996, 150–161.
- [8] Hatchuel A., Le Masson P. and Weil B. The development of science-based products: managing by design spaces. *Creativity and innovation management*, 14, 4, 2005, 345-354.
- [9] Inselberg A., *Parallel coordinates*. Springer US, 2009.
- [10] Keim D., Kohlhammer J., Ellis G. and Mansmann F. (eds.) *Mastering the Information Age: Solving Problems with Visual Analytics*. Eurographics, Goslar, 2010.
- [11] Maher M.L. and Poon J., Co-evolution of the fitness function and design solution for design exploration. *Proc. IEEE International Conference Evolutionary Computing*, 1995, 240-244.
- [12] Mayorga A. and Gleicher M., Splatterplots: Overcoming overdraw in scatter plots. *IEEE transactions on visualization and computer graphics* 19.9, 2013: 1526-1538.
- [13] Miles J.C., Sisk G. and Moore C.J., The conceptual design of commercial buildings using a genetic algorithm. *Computers & Structures* 79, 2001, 1583-1592.
- [14] Monks M., Oh B. and Dorsey J., Audiooptimization: Goal based acoustic design. *Computer Graphics and Applications, IEEE Journal* 20, 3, 2000, 76–90.
- [15] Mourshed M., Shikder S. and Price A., ϕ -array: a novel method for fitness visualization and decision making in evolutionary design optimization. *Journal of Advanced Engineering Informatics* 25, 2011, 676–687.
- [16] Sosa R. and Gero J.S., The creative value of bad ideas, in: *Proc. CAADRIA 2013*.
- [17] Turrin M., von Buelow P. and Stouffs R., Design explorations of performance driven geometry in architectural design using parametric modeling & genetic algorithms. *AEI* 25, 4, 2011, 656-675.
- [18] Turrin M., von Buelow P., Kilian A. and Stouffs, R., Performative skins for passive climatic comfort: a parametric design process. *Automation in Construction* 22, 2011, 36-50.
- [19] Unwin A., Theus M. and Hofmann H., *Graphics of Large Datasets: Visualizing a Million*. Springer Science & Business Media, 2006.
- [20] Woodbury R. and Burrow, A.L., Whither design space? *AIEDAM* 20, 2, 2006, 63–82.
- [21] Woodruff M.J., Reed P.M. and Simpson T.W. Many objective visual analytics: Rethinking the design of complex engineered systems. *Struct Multidisc Optim* 48, 2013, 201-219.
- [22] von Buelow P., ParaGen: Performative exploration of generative systems. *J. IASS* 53, 4, 2012, 271-284.
- [23] von Buelow P., Improving generative design through selective breeding, in: *Proc. IASS-SLTE*, 2013.
- [24] von Buelow P., Using database storage to improve explorative optimization of form critical structures, in: *Proc. IASS-SLTE*, 2014.