Strategies for the multivariate representation of micro-environmental design space. Considerations on parallel coordinates plot

Brunetti, Gian Luca [1]

Keywords: parallel coordinates plots; multi-objective analysis; parametric analysis; environmental simulation

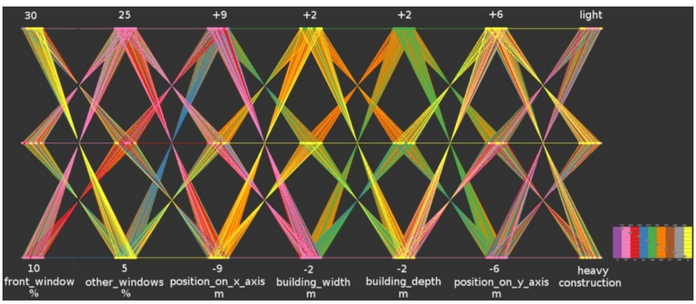
Abstract This paper reports the results of experiences with the use of parallel coordinates plots for the analysis of multi-objective, multivariate problems for architectural design at the micro-environmental scale, as possible through the support of thermal and lighting simulation analyses.

Parametric analyses applied at the micro-urban level may produce a wealth of information which is difficult to exploit for what both design and decision-taking are concerned. One reason for this is the sheer difficulty of producing convincing visual representations of the complex relations between the most relevant variables in play.

Pragmatic strategies for the use of parallel coordinate plots have been experimented by the author to produce pictorially eloquent representations blending quantitative and quantitative information. This was mainly done through three strategies:

- texturing, allowing for an increase of understandability of large data domains;
- layering, through the superimposition of suitably treated representations of distinct objective functions;
- filtering, through the application of threshold levels to objective functions.

Fig. 1 Parallel coordinate plots in discontinuous colours regarding the resultant maximum temperatures in a building obtained with parametric simulations through variations of the position, shape and orientation of a building on a site, as shown in Fig. 3. In clear nuances (yellow) are the most advantageous performances. The instances producing mean daylight factors lower than 3% in a room horizontal plane have been filtered out from the representation. Among other things, the following situations may be noted, from left to right: an advantage of orientations depending from other factors; an advantage of the largest building widths; an advantage of the largest building depths; an advantage of warping angles depending on other factors; an advantage of the smallest front window size; an advantage of the smallest other windows' size; an advantage of the heaviest construction solution



1. Research Unit for Quality and Sustainability of the Built Environment, Dipartimento di Architettura e Studi Urbani (DAStU), Politecnico di Milano, Milan, Italy

Introduction

The full potential of multi-dimensional analyses for designers is most often best exploited through visual representation methods. Therefore the adoption of suitable representation methods may be a key success factor for procedures of the kind here described, second only to the sheer availability of results.

Statistical representation of variables in building design is usually done by 2D or 3D Cartesian representations. But even in a 3D Cartesian representation, when using colours to convey added information, only four variables can be clearly represented at once. Five at best, in a still image, when also symbols are used. And multidimensional representations (in the sense of more than 3D) in Cartesian space seem today too counterintuitive to be considered for applications in design. Therefore, more than one distinct graph have to be used to convey information for a typical multi-dimensional inquiry aimed to the early stages of design when Cartesian representations are used, at the risk of missing important relations between objective functions and variables. For this reason, it is opinion of the author that a quali-quantitative (qualitative and quantitative) use of parallel coordinates representations may be a very useful procedural resource for building design.

The approach that will be adopted in the paper will be to represent the multi-dimensional content produced by simulation-based parametric exploration by the means of parallel coordinate plots, by adopting devices to make them as understandable and operatively useful as possible.

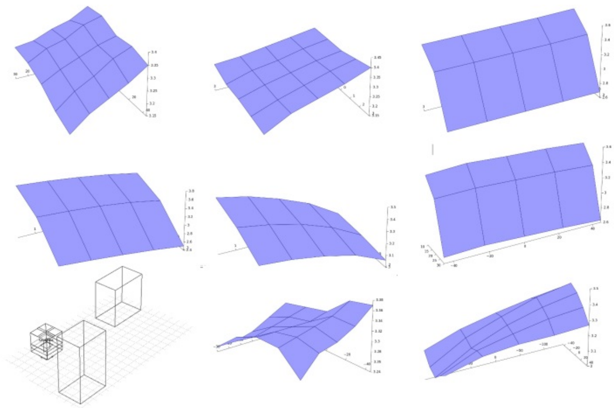


Fig. 2 An example of 3D Cartesian representation for a case study of a free-standing greenhouse. In the lower left corner: the model in its context, with solar obstructions. The variables taken into account were resultant mean temperatures, x and y position on plane, azimuth variation, shape ratio, and glazings type.

Background about data representations for parametric analysis aimed to building design

Simulation applications to foresee and model environmental conditions are today often used in the context of parametric analyses. Parametric analyses often integrate certain kind of optimization search system aiming to reduce the search size explosion due to the exponential relation linking the number of parameters in play and search size in linear search systems.

So called parametric CAD tools share the term "parametric" with reference to the parametric modelling allowed by the object-oriented programming technologies they are developed with; but they only seldom and partially are conceived to support the kind of analyses of the outcomes of parameters' variations adopted in the prevailing morphing-supported optimization approaches.

An advantage of theirs is that they are coded like objects and that therefore make possible to attach dynamic properties to models. But the other side of the coin is that they often produce complexity for the sake of CAD functionality, at the expense of morphing agility.

Interesting experiences made with animation morphing tools to morph models for thermal simulations had to overcome the complexity of closed-source APIs of commercial CAD tools (Franconi, 2011).

The author has prototyped a parametric morpher, named OPTS (Brunetti, 2008), targeting the ESP-r simulation platform to allow object-like properties to the models by implementing the strategy known in Artificial Intelligence as propagation of constraints, or constraint propagation.

Parametric analyses often requires the aid of multivariate representation methods, due to the broad spectrum of data they can produce and to their static nature. This requires that, in order to have a complete picture of the design options and their consequences, their representation has to be both deep and broad.

An opposite approach is that of real-time simulation, which is aimed to give the designer a direct feeling of how models respond to change (Clarke, 1985-2001), reducing the scope of broad representations. This is due to the fact that if the models have dynamic properties, the need of representing their behaviour is reduced, since it can been observed through real-time changes.

The parametric data obtained through simulation programs are usually multi-objective and suitable to be treated by derivative-free optimization methods, of which the various types of pattern search are the most used. This is because the calculations of simulation programs are usually dealt with by complex layering of conditions which make the results unsmooth.

This is why curve-fitting of objective functions obtained by simulations is often necessary to use the results through derivative-based optimization methods.

There is at least one case – BuildOpt - in which simulation applications have been conceived based on smooth models (Wetter, 2005), to ease the adoption of derivative-based optimization methods, in response of the fact that the data produced by simulation-based parametric analyses are usually so discontinuos and unsmooth at mathematical level, that have to be treated as if coming from so-called black boxes (tools producing responses that have to be accepted without understating how they were produced). It is not a case that both the well known optimization applications Genopts (Wetter, 2000) and JEPlus (Zhang, 2011) are principally based on derivative-free pattern search algorithms.

General background and issues about parallel coordinates plots

Parallel coordinate plots have been invented by Maurice d'Ocagne in 1885 (D'Ocagne, 1885) and re-discovered by Alfred Inselberg in 1959, who backed their diffusion (Inselberg, 1985; Inselberg, 2009) until today.

In last three decades the interest for them have gradually gathered momentum, also because multidimensional representations have become less and less expensive to obtain, and, parallel coordinate representations have been used in many fields, including environmental analysis.

Parallel coordinate representations are distinguished by the fact that the variables' values are plotted on a coordinate system formed by parallel segments, one for each considered parameters, and the instances (vector components) to be plotted are represented as lines connecting points on the segments constituting that coordinate system, which may express the value of objective functions through colors.

That way, the behavior of a variable may be represented through colours, or by position on the coordinate systems, or compounding the two strategies, which may allow to stress the importance of that variable other the other ones. Moreover, a variable can be expressed in 3D, plotting a parallel coordinate representation with a "z" depth, seen in perspective or in parallel projection. This adds one more means of representation of objective functions in a plot, which is especially suited to bi-objective problems, where one objective function may be expressed through colors and the other through depth (z coordinate) in 3D space.

In theory, this wealth of options and flexibility in representing variables corresponds to a great power to convey nuances of meaning in representation without limiting the possibility to front complexity. In practice, this is mostly true only for small or medium data sets. This is due to a well-known difficulty in spotting the individual instances of variables with parallel coordinates when they are plotted in high numbers. This raises important questions, because the kinds of multivariate and multi-objective problems which are suitable to be represented with parallel coordinates plots are often also characterized by high numbers of variable instances. This is why a great deal of techniques and representation strategies for parallel coordinates plots are aimed to improve the clarity and readability of single vector components in the maze of complex bulk of data.

A useful strategy of these is named “texturing”. All parallel coordinate representations in this paper are obtained through texturing. Texturing is especially suited to parametrically simulated datasets. It uses series points distributed on small segments perpendicular to the parallel coordinate lines instead of single points to connect the segments describing variable values, which allows the data to be readable in a sort of clearer, pictorial, designer-friendly manner, as showed in Fig. 3. The same reason – the sake of clarity - is also why the 3D information component of parallel coordinate plots, when present, is often used to strengthen the message conveyed by colours, in a sort of “unison” representation (Fig. 8).

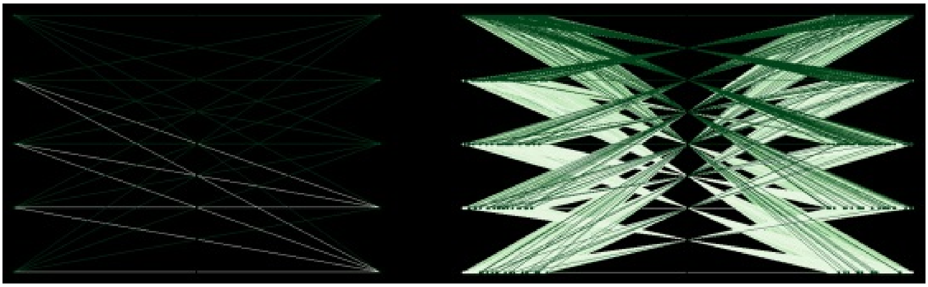


Fig. 3 Comparison between an ordinary coordinate plot representation (left) and a textured coordinate plot representation (right) from the same dataset

Experimental considerations of parallel coordinate plots for multivariate and multi-objective problems, and conclusions

Considerations will be here be made based on test-cases defined by the author with the aid of the cited application named OPTS as a morpher (Brunetti, 2013), to parametrically manage the well-known, ESP-r advanced energy simulation coupled with the lighting simulation program Radiance (Ward, 1994).

In all tested cases, the objective was the placement of an architectural object on a lot. A micro-environment was modelled by the means of solar and wind obstructions (evaluated through a pressure coefficient-based ventilation nets), updated at any change in the scene. This approach has allowed the automated production of broad sceneries of variables.

In general terms, important advantages which have been verified through use for parallel coordinate representations are that they are fully multidimensional, with no limit in the number of variables, and that they are very concise, suitable to deliver a great quantity of information simultaneously.

More specific considerations stemming from the tested cases are the following.

1. An advantage of parallel coordinate representations is that they make possible to obtain clear distinctions of the ranges of performance, which is very useful when representations have to support unambiguous decisions. This is possible by assigning a different colour to each level of performance (objective function), obtaining qualitative representations emphasizing the top- and bottom-performing combinations of variables' values. This possibility is very useful, for example, to highlight top-performing and bottom-performing combinations of variables' values at the expense of other values, which is a common objective in analyses. (see Fig. 1).

2. Parallel coordinates plots make possible to mix quantitative and qualitative visual criteria in order to emphasize particular aspects of performances. In particular, when a range of nuances of a certain homogeneous colour is assigned to a certain performance indicator (objective function - for instance, one whose increase produces an advantage: utility function) a more telling representation than the former from a scientific point of view may be obtained, more intuitively conveying information about performance though colours (see Fig. 5 and 6).

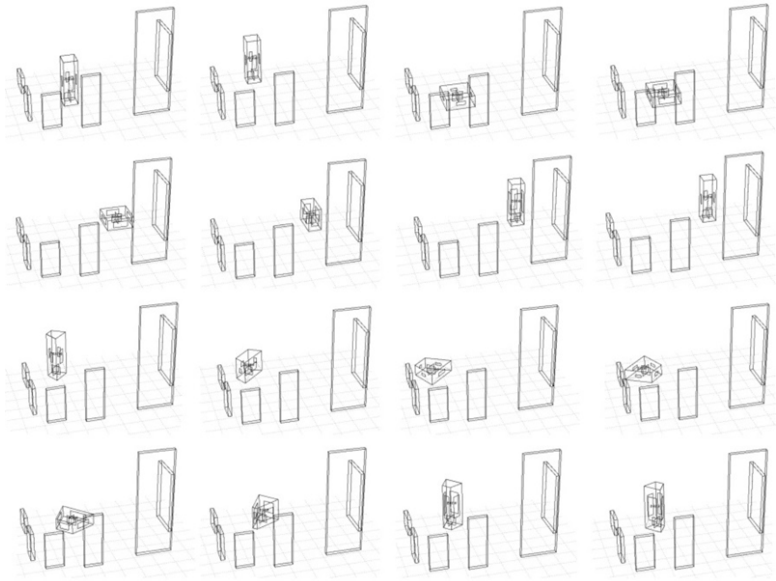


Fig. 4 Morphing steps for a building in a square modelled with solar and wind obstructions. The changing shape is that of the building, and the fixed ones are the obstructions. The variations regarded position, shape and construction.

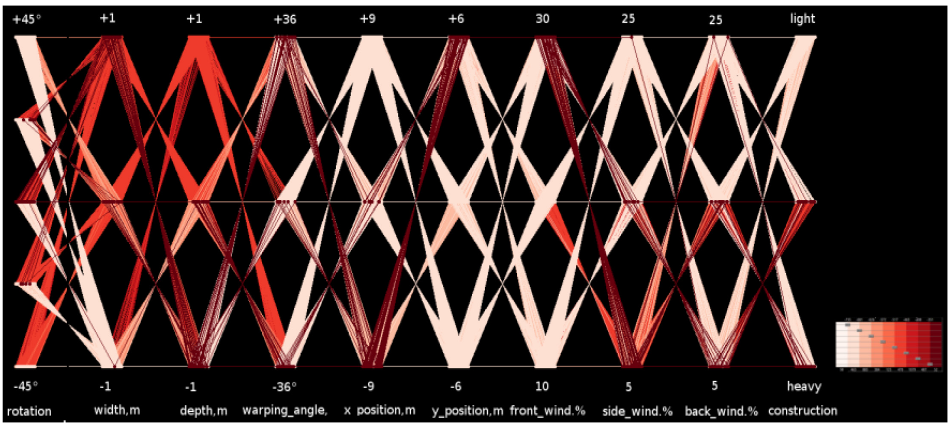


Fig. 5 Parallel coordinates plot in one colour representing heating loads in February, derived from the parametric analysis showed in Fig. 4. Of the 98415 simulated instances, the 2000 top and the 2000 bottom performing ones are shown. The darker the colour, the better the performance (smaller loads). Here the following situations may be noted: an advantage of orientations toward south or 22.5° south-west; and advantage of large widths combined with medium or shallow depths; most advantageous warping angles depending from other conditions; an advantage of western and northern positions of the lot; and advantage of the largest front window size and of the smallest other windows' size; and an advantage of heavy and medium weight construction solutions.

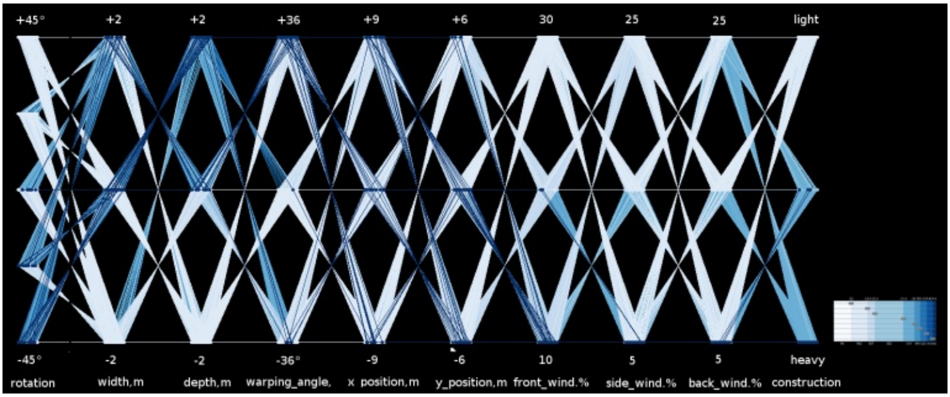


Fig. 6 Parallel coordinates plot in one continuous colour regarding the same case study of the graph above and representing the mean maximum resultant temperatures in August. The darker the colour, the better the performance (lower temperatures). Here the following situations may be noted: an advantage of orientations toward south-east and south; and advantage of large widths combined with large depths (producing large bases and small heights); an advantage of warping angles bringing convexity to the front façade; and advantage of positions on the east-west axis depending from other factors; an advantage of southern, obstruction-shaded positions; and advantage of smaller windows and heavy construction solutions.

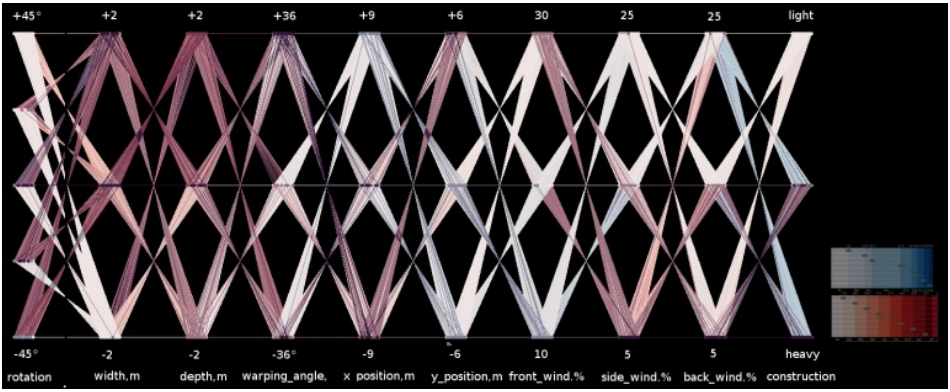


Fig. 7 Quali-quantitative parallel coordinates plot in which thermal loads and maximum temperatures from the previous two graphs have been superimposed. The nuances are telling about intermediate situations. Pale reds signal middle-high winter loads coupled with high summer maximum temperatures, while pale blues signal middle-low winter loads coupled with low summer maximum temperatures.

In that situation, it is possible to superimpose in transparency the representation of a second objective function to the first one, plotted in a continuous colour complementary to the first, so to convey additional information about the balance of the two. This may especially be useful where there are two objective functions which do not work in synergy, but anti-synergetically: typically, a utility function and a cost function. The superimposition of the two functions represented in complementary colours allows to fuzzily distinguish the four possible main different groups of combinations of results (see Fig. 7).

3. Thanks to the possibility of representing broad ranges of variables with homogeneous criteria, parallel coordinates representations encourage to blur the distinction between model's data (variables) and results (objective functions). In doing so, they may ease the operation of

treating model's generative data (among them, typically, geometrical and construction data, for instance) and objective functions (performance indicators) on the same plan, with the same right; each one, in possibility, an incognitum or a datum. Several heterogeneous and useful performance indicators may coexist on a graph, together with the chosen principal one or two, and may be indicated by colours, and/or be used to filter the results. This may be, for example, the role of objective functions regarding light - like daylight factors or mean lighting coefficients - in a design tested on the basis of thermal behaviour. Thermal loads and temperatures then may be took as the main objective functions, and daylight data may be used as filtering conditions and/or superimposed conditions. This may be for any reasonable performance indicator, both heterogeneous and homogeneous to the principal ones. In Fig. 1, for instance, daylight factors have been used as a filtering condition (and added objective function) for the representation of an otherwise bi-objective function.

In other words, parallel-coordinates-based representations encourage to treat simple or compound objective functions (like, for example, thermal comfort levels, lighting comfort levels, illuminance levels, or acoustical reverberation time – which are typically considered results) like any other kind of model's descriptors; like variables; in the same conceptual domain of non-performance-related qualities belonging to a model. This may encourage dealing with illuminance levels, for instance, as a major design constraint, like volume, instead of a major constraints' result.

The result of the design procedures were to embody the information about microclimate conditions on the building site into the results of exploration operations based on variations on a model (a meta-model, not necessarily nor quantitatively nor qualitatively similar to the to-be-designed thing) used as an indirect measurement tools of the suitability (the degree of aptness) of the examined types of solutions. The so-obtained multi-dimensional representations have been of use to ease the evaluation of multiple objectives, as especially required in early design.

4. Data produced in the context of parametric simulation are rarely suitable to untreated parallel coordinate plots. They are usually suited to "textured" plots. This is because a good part of the data to plot following parametric simulations are not those of results, but those showing parameters' variations. Which, unless randomized, are regular, and usually covering just a few discrete values. If a "texturing" technique were not adopted in those cases, a great deal of the available data would therefore be made invisible, due to overlayering. It must be noted, however, that in spite of texturing, a maximum of about 5000 instances was the upper limit verified by the author to obtain an acceptably clear representation for the problems which have been examined.

5. 3D parallel coordinates representations can even be produced through ordinary CAD programs (Fig. 8), if appropriate scripting techniques are adopted. But for a greater clarity, further added devices could be probably of aid, or maybe even mandatory.

Acknowdedgements

Parallel coordinates plots have been used by the author in the research "Systemic integration of renewable energy sources technologies in the built environment" funded by the Italian Ministry and developed during the years 2011 and 2012 to evaluate the suitability of building sites solar systems, whose coordinator was prof. Gianni Scudo, Politecnico di Milano. The 2D coordinate representations presented in this paper were obtained with the open source tool *Ggobi* (Swayne, 2003).

References

Brunetti, G.L.(2008). A Design Optimization Tool for Building Energy Simulation, *WREC* X Conference Proceedings, Glasgow, p. 781-784.

Brunetti, G.L. (2013). Procedural meta-models for architectural design praxis. *Techne* ; 5:170-176.

Clarke, J. (1985). *Energy Simulation in Building Design*, Butterworth Einemann, London 1985-2001.

D'Ocagne M. (1885). *Coordonnées parallèles et axiales. Méthode de transformation géométrique et procédé nouveau de calcul graphique déduits de la considération des coordonnées parallèles*. Paris, Gauthier-Villars.

Franconi, E. A (2011). Method for the design and analysis of parametric building energy model, *IBSPA* Conference, Sidney.

Inselberg A. (1985) The Plane with Parallel Coordinates. *Visual Computer*. 1(4):69–91.

Inselberg A. (2009). *Parallel Coordinates: Visual Multidimensional Geometry and its Applications*. Springer, New York.

Swayne, D.F., Temple Lang, D., Buja, A., Cook, D. G., (2003) Gobi: Evolving from XGobi into an Extensible Framework for Interactive Data Visualization, *Computational Statistics & Data Analysis*. 2003; 43:423-444.

Ward, G.J. (1994). The Radiance Lighting Simulation and Rendering System, Computer Graphics (Proceedings of '94 SIGGRAPH conference). p. 459-72.

Wetter, M. (2000). Design Optimization with GenOpt, Research Report, Simulation Research Group, Laerence Berkeley Laboratory.

Wetter, M. (2005). BuildOpt—a new building energy simulation program that is built on smooth models. *Building and Environment*. 40(8):1085–1092.

Zhang, Y. (2011-2012). *JEPlus User's Manual*. Web document, http://www.iesd.dmu.ac.uk/~jeplus/wiki/doku.php?id=docs:manual_1_3

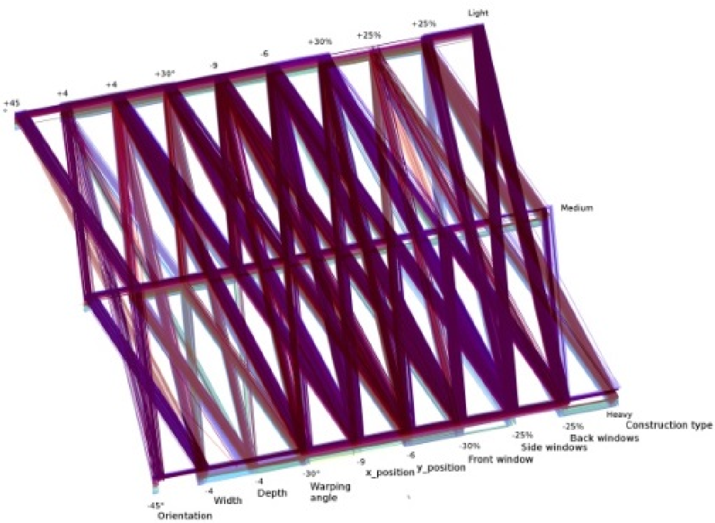


Fig. 8 3D parallel coordinates plot in two colors obtained by programming a conventional CAD program.