Selection criteria for energy performance improvement measures in traditionally constructed non-domestic buildings

Citation for published version:
Banfill, PFG & Strachan, ME 2014, Selection criteria for energy performance improvement measures in traditionally constructed non-domestic buildings. in *Proceedings of the 8th International Conference on Improving Energy Efficiency in Commercial Buildings, 2014.* European Commission Joint Research Centre/Institute for Energy, Brussels, pp. 293-301, Improving energy efficiency in commercial buildings, Frankfurt, Germany, 1/04/14.

Published In:
Proceedings of the 8th International Conference on Improving Energy Efficiency in Commercial Buildings, 2014

Selection criteria for energy performance improvement measures in traditionally constructed non-domestic buildings

M.E. Strachan, P.F.G. Banfill

CO2 Estates, Alloa, Scotland, Heriot-Watt University, Scotland

Abstract

In the context of the energy-led refurbishment of traditionally-constructed non-domestic buildings, a two-stage Delphi exercise has been carried out in order to identify the criteria that professionals consider to be important in the selection of energy performance improvement measures, and to establish their relative importance. Based on the opinions of a panel of experts, 22 assessment criteria were identified, covering short-term outcomes, long-term issues, end of life issues, and whole life impacts. A paired comparison survey enabled the weightings of relative importance for the criteria to be determined. Measures fell into one of three broad groups, expressed on a scale of 0-100, such that the sum of the weights of all 22 measures was 100. Measures of relatively high importance (8±1%) included capital cost, potential energy and carbon savings, financial payback and impact on the building's vapour permeability. Measures of medium importance (4.5±1%) were impact in internal air movement, loss of significant original building fabric, impact on internal layout, appearance and occupant comfort, environmental impact and availability of grants or subsidies. Eight further measures were ranked of low importance (2-3%).

Introduction

In the UK, about 40% of energy use and CO2 emissions is associated with existing buildings, so significant reductions can only be achieved by refurbishment. This paper deals with traditionally-constructed non-domestic buildings which are common in the building portfolio of service sector and public sector organizations, providing them with a presence in many major city centres. The professionals responsible for making decisions on the measures to be adopted in order to improve their operational performance are faced with a plethora of technical and other information, which must be weighed carefully against the needs and aspirations of their corporate employers. Following a review of the literature on the decision support tools available for such professionals, Strachan and Banfill [1] developed a seven-step generic decision support tool for energy led refurbishment of non-domestic buildings, which could be used by property managers who are responsible for a non-domestic building or a building portfolio. The seven steps are:

1. Building assessment to gain a holistic view of performance and record it in a database.

2. Identification and selection of energy demand interventions, such as change in occupancy culture and behaviour, changes in lighting and appliances, and changes to building fabric, services equipment and controls.

3. Simulation of building energy performance after application of energy performance improvement measures (EPIMs) to reduce energy demand.

4. Identification and selection of energy supply interventions, through adoption of low carbon / renewable energy sources.

5. Simulation of building energy performance after application of both demand and supply interventions.

6. Energy management action plan to advise the user on how to manage their improved building.

7. Continuous improvement through periodical review of performance.

The objective of the work described in this paper is to develop step 2 into a practical tool for the property professional responsible for an existing office building of traditional construction, defined as pre-1919, mass masonry construction, originally single glazed, originally without insulation materials built into the fabric, and likely to have high air infiltration levels. These professionals are faced with a range of EPIM alternatives in a complex decision making process that can be classed as Multiple Attribute Decision Making (MADM). MADM is characterised by alternatives, numerous attributes, attribute weights and numerical values in inconsistent units that are difficult to compare [2]. In this study the alternatives are the EPIMs, which have multiple attributes that can be assessed both qualitatively and quantitatively. From this mixture, MADM involves the evaluation of information and prioritisation of solutions, which in turn requires some form of weighting of the various attributes. Since individual decision makers are likely to have different opinions on the relative importance of each criterion, it is important to develop a consensus. This was done using a two-stage Delphi process which, first, established an agreed set of assessment criteria and, second, developed weightings for each criterion.

The Delphi technique

Delphi is a method that supports comprehensive decision making, planning and problem solving, and has been used in a wide range of sectors – construction, education, healthcare, information technology, marketing and transport [3-8]. The core of the method is the obtaining of statistically valid consensus between a group of experts in a specific field, based on their knowledge and experience, implemented through a series of iterative questionnaires, combined with controlled, anonymous feedback [9]. The strength of Delphi is that it creates an environment where each expert can think independently, without the pressure of a group scenario, where forceful personalities can unduly influence the thinking of the group [10] and even cause direct confrontation [11]. An appropriate Delphi expert is one that is a highly skilled specialist in the subject [12] but is also open to revision of their views when presented with new information [13]. Between ten and fifty experts are considered suitable for Delphi [14], so recruitment of experts can be challenging, but a greater challenge can be the time taken within and between each survey round to allow the experts to consider the content of each round of questions and between each survey round for the researcher / facilitator to analyse and draw up feedback of the expert views [11].

Weighting multiple assessment criteria

Establishing the relative importance of a set of assessment criteria that have been drawn up in a Delphi process is complex because of the conflicting nature of the criteria. Among available methods Bartlett [15] reviewed fixed point scoring, rating, ordinal ranking, graphical weighting and paired comparisons. Since methods requiring participants to distribute points over several items at once are too demanding [16] the paired comparison method has the advantage of asking for only two criteria to be considered at once: participants compare their relative importance in a systematic questionnaire. The comparison is facilitated by a numerical scale but the questionnaire can be lengthy because every criterion must be assessed against every other one in the set and the comparisons must be carefully organized to minimize respondent fatigue.

Methods

A panel of 13 experts was recruited to take part in a two-stage Delphi process, each with specialist experience within one or more of five relevant sectors:

(i) Client. Experts who work within an organisation with a significant building portfolio, who are involved in the management and works to that portfolio.

(ii) Guidance. Experts who work within an organisation who set standards or guidelines for construction standards or experts who are involved in and promote research into energy in buildings.

(iii) Heritage. Experts working within an organisation who work to safeguard the historic built environment.

(iv) Industry. Experts who work within the construction industry.

(v) Non-heritage. Experts who have a technical background with an understanding of the historic built environment but are involved in a broader range of building types.

Delphi stage 1

In stage 1, the experts developed an agreed set of criteria for the use of built environment professionals in assessing the suitability of an EPIM for an existing building. This involved three rounds of online questionnaires, administered by Surveymonkey®. In round one, each participant received a questionnaire consisting of an initial list of 15 criteria, with seven questions, some in yes/no format and some requiring a comment (table 1). Their replies, elicited over two further rounds, led to an agreed list of 22 criteria, which were then used in stage 2. Experts were given two weeks to reply to each round of questionnaires, in line with recommended practice [11].

Table 1 Delphi survey questions

Number	Question	Response format
1	In your opinion, does the list contain sufficient criteria to assess the suitability of an Energy Performance Improvement Measure?	Yes/No
2	Are there any criteria that should be added to the above list?	Yes/No
3	If yes, then please describe what additional criteria should be added.	Comment box
4	If yes, then please explain why these additional criteria should be added.	Comment box
5	Are there any assessment criteria that should be omitted from the above list?	Yes/No
6	If yes, then please describe what criteria should be omitted.	Comment box
7	If yes, then please explain why these criteria should be omitted.	Comment box

Delphi stage 2

In stage 2, the agreed assessment criteria were weighted in terms of their relative importance in the decision-making process, using a paired comparison questionnaire. In order to compare every criterion against every other criterion, the number of paired comparisons *N* is given by the formula:

$$N = m\ (m-1)\ /\ 2$$

where *m* is the number of criteria. The 22 assessment criteria therefore generated 231 pairs of comparisons, which were presented to each respondent as a nine-point Likert scale (table 2).

Table 2 Paired comparison questionnaire (respondents tick the chosen box) and scores assigned

		Very strongly more important	Strongly more important	Moderately more important	Slightly more important	Equally important	Slightly less important	Moderately less important	Strongly less important	Very strongly less important	
	Criterion A	☐	☐	☐	☐	☐	☐	☐	☐	☐	Criterion B
Bipolar scoring		+4	+3	+2	+1	0	-1	-2	-3	-4	
Unipolar scoring		9	8	7	6	5	4	3	2	1	

Experts were asked to compare criteria A and B and tick the appropriate box. The comparisons were presented to the experts in random order to maintain their engagement in the exercise. In view of its size, the questionnaire was administered on paper, formatted with the help of Excel, and posted to the experts. Experts were asked to reply in 28 days but this period lengthened because of the number of questions asked but the whole process was completed within the 45 days recommended practice [11]. The results were analysed in three ways – (i) negative to positive bipolar ranked assessment, (ii) bipolar sum of differences, and (iii) unipolar ranked assessment.

Results and analysis

All 13 experts participated in stage 1 but, presumably as a result of the lengthy questionnaire, only 11 participated in stage 2. Nevertheless, this number exceeds the minimum of 10 required for statistical significance [14] and the weightings presented below are therefore considered to be valid.

In stage 1 the experts' replies expanded the list of 15 initial criteria to 23 in round two, whereupon the adjusted list of criteria was administered again, using the same seven questions. In the third and final round, in response to the experts' suggestions the criteria were collected into three categories based on the life cycle stage(s) – installation, operation and end of life - to which the criteria relate. Additionally, two criteria two were merged, leading to the final 22 criteria (table 3 and 4).

Table 3 EPIM Assessment Criteria categorized according to their relevant life stage

Energy Performance Improvement Measure (EPIM) Assessment Criteria (AC)					
AC result realised in the short term (beginning of EPIM's useful life)		AC result realised in the long term (i.e. during / end of EPIM's useful life)			
EPIM Installation		EPIM Operation		EPIM Disposal	
1	Capital cost	9	Potential energy/carbon savings	20	Disposal cost of EPIM at end of useful life
2	Availability of grants, tax allowances and other financial incentives	10	Financial payback		
3	Ease of installation of EPIM	11	Change to maintenance costs		
4	Loss of significant, original building fabric	12	Ease of maintenance of EPIM		
5	Requirement of planning and/or building control approvals	13	Reliability of EPIM's performance		
6	Level of disruption to building occupants during works	14	Degradation of EPIM's performance		
7	Impact on building's appearance	15	Training building occupants in the use of new system(s)		
8	Impact on building's internal space/layout	16	Level of improvement in building occupants' comfort		
		17	Impact on existing building services		
		18	Impact on building's internal air movement/ ventilation		
		19	Impact on building's vapour permeability/ breathability		
21	Embodied energy/carbon of EPIM				
22	Environmental impact of EPIM				

Table 4 Definitions of EPIM assessment criteria

No.	Assessment criterion	Definition
1	Capital cost	Initial cost incurred to purchase the EPIM, including all associated transport, labour and materials.
2	Availability of grants, tax allowances and other financial incentives	The availability of financial incentives for the implementation of particular EPIM's.
3	Ease of installation of EPIM	Also known as 'buildability'. The level of difficulty associated with the installation of an EPIM, including ease of transport to and movement on site.
4	Loss of significant original building fabric	Some EPIM's installation will have a low visual impact but may result in loss of significant, original building fabric.
5	Requirement of planning and/or building control approvals	The likelihood of requiring some form of formal approval for the installation of an EPIM, including Listed Building Consent where applicable.
6	Level of disruption to building occupants during works	The level of disruption caused by the installation of an EPIM on the building occupants' working environment, and consequently their productivity.
7	Impact on building's appearance	The impact the installation of an EPIM will have upon a building's appearance, both externally and internally.
8	Impact on building's internal space/layout	The installation of some EPIM's could impact upon the gross internal floor area or the internal layout of the building.
9	Potential energy/carbon savings	A quantitative measure of the energy savings and associated carbon emission savings of installing an EPIM.
10	Financial payback	A measure of the time required to recover the initial cost invested.
11	Change to maintenance costs	A potential increase or decrease in the building user's maintenance budget due to the installation of an EPIM.
12	Ease of maintenance of EPIM	The level of difficulty associated with the maintenance of an EPIM and any associated equipment or materials. Including the availability of spare parts over the lifetime of the EPIM.
13	Reliability of EPIM's performance	The reliability of an EPIM's performance. Risk of failure in meeting predicted energy savings, as well as any other performance criteria.
14	Degradation of EPIM's performance	The potential year on year reduction in the EPIM's ability to deliver energy savings.
15	Training building occupants in the use of new system(s) post refurbishment	The level of training and regular re-training required of building occupants to ensure the EPIM is operated at its maximum efficiency.
16	Level of improvement in building occupants' comfort	The level of improvement in indoor environmental quality due to EPIM installation, consequently improving the building occupants' comfort levels and potentially, worker productivity.
17	Impact on existing building services	The impact the EPIM's installation will have upon the existing building services (BS), including building fabric improvements, as these will change the internal environment and how it interacts with the BS. Some BS-related EPIM's can have a negative impact on the existing plant and its maintenance, and this must be considered.
18	Impact on building's internal air movement/ventilation	The impact of the EPIM's installation on how the existing building deals with air movement. A negative impact could lead to serious air quality and condensation issues. Also, whether changes to the building's ventilation strategy need to be considered as a result of this EPIM.

19	Impact on building's vapour permeability/'breathability	A qualitative measure of the impact an EPIM's installation has on the building fabric and how it interacts with moisture. Whether or not that EPIM is compatible with the existing construction form.
20	Disposal cost of EPIM at end of useful life	The financial cost of removing and disposing of the EPIM and any associated parts at the end of their useful life.
21	Embodied energy/carbon of EPIM	The total energy/carbon inputs required to manufacture an EPIM and its associated materials, from extraction of raw materials to reuse/recycle/disposal. This also covers the issue of EPIM availability, in terms of the energy/carbon cost of sourcing and transport.
22	Environmental impact of EPIM	The level of pollutants/environmental cost accumulated in the manufacture of an EPIM and its associated materials, from extraction of raw materials to reuse/recycle/disposal.

The results of the paired comparison questionnaire were analysed in three ways. In the negative to positive bipolar ranked assessment, the 9 point scale (table 2) essentially represents two opposing poles with a central point denoting equal importance. The scores obtained by a criterion in each of 21 comparisons with the other criteria are summed and the total score can therefore range from +84, denoting an overwhelmingly important criterion, to -84, denoting the reverse. In practice, most but not all of this range was used. In the bipolar sum of differences analysis, the criterion of lesser importance in each pair is scored zero instead of the negative value. This avoids over-emphasising the negative view of the less important criterion, and truncates the total score range to 0 to +84. The option of equal importance remains, with both criteria scoring zero. This analysis is considered to be the most appropriate for this investigation. Finally, the unipolar ranked assessment ascribes one pole in each pair as being more important, and can be viewed as a less natural fit with the paired comparison approach, despite having been used in this way previously [17]. In this case the total score can range from +21 to +189.

In each analysis, the total score gained by each criterion was averaged over all the 11 experts, and the mean values normalised to enable them to be presented as weightings in percentages. Table 5 shows the resulting scores for each criterion, and it is clear that the top five criteria in each case are the same and appear in the same order, and that the unipolar analysis gives a narrower distribution of values than the two bipolar methods.

As can be clearly seen in Figure 1, the criteria fall into three clear groups according to their weightings – a group of low importance assigned 2-3%, a group of medium importance assigned 4.5±1% and a group of high importance assigned 8±1%.

Discussion

The assessment criteria are largely self-explanatory but some comments may be helpful. It can be seen from tables 3 and 4 that the criteria for each EPIM are framed in neutral terms, such as ‘impact on …’ or ‘level of …’, to avoid prejudicing the experts’ comparisons. However, since the objective of an EPIM is to save energy and reduce carbon emissions criterion 9 was framed explicitly as ‘potential energy / carbon savings’. The word ‘significant’ is included in criterion 4 because it discourages the user from being too cautious and rejecting a beneficial EPIM because of the loss of fabric that would not be considered important in heritage conservation terms. In some cases, separate criteria were merged when it became obvious that, for example, payback period and financial savings are essentially equivalent, leading to criterion 10 ‘financial payback’, but are, however, unconnected with maintenance costs, thus retaining criterion 11 ‘change to maintenance costs’.

Figure 1 shows that three criteria - capital cost, potential energy / carbon savings and financial payback - are assigned the highest importance within the set. Since the objective of any energy improvement measure is to save energy and reduce carbon emissions within a business context, it is not surprising that these are the top three criteria. It confirms the experts’ pragmatic view that the typical user of the decision support tool is concerned with the financial impact of their decision on their organization. However, the fact that there are a further 19 criteria, together accounting for over 70% of the weighting scores, shows that they believe that decision making should be informed by much more than these three fundamental factors.

Table 5 Comparison of weightings (%) of the three analyses

Number	Criterion	Bipolar negative	Bipolar sum of differences	Unipolar
1	Capital cost	9.03	9.30	5.66
2	Availability of grant, etc	3.61	3.64	4.31
3	Ease of installation	2.50	2.42	4.04
4	Loss of significant building fabric	5.21	5.35	4.71
5	Requires planning approvals	1.83	2.56	3.87
6	Level of disruption to occupants	2.27	2.82	3.98
7	Impact on appearance	4.80	4.93	4.61
8	Impact on internal layout /space	4.93	5.22	4.64
9	Potential energy savings	8.79	8.77	5.60
10	Financial payback	8.31	8.22	5.48
11	Change in maintenance costs	2.86	2.87	4.13
12	Ease of maintenance of EPIM	2.53	2.27	4.05
13	Reliability of EPIM	5.67	4.85	4.82
14	Degradation of EPIM	4.48	3.77	4.53
15	Training of occupants in new systems	0.95	2.82	3.66
16	Level of improvement in comfort	4.55	4.48	4.55
17	Impact on existing building services	2.89	2.48	4.14
18	Impact on internal air movement	6.09	5.45	4.93
19	Impact on vapour permeability	7.49	7.19	5.27
20	Disposal cost of EPIM	2.00	2.08	3.92
21	Embodied energy of EPIM	4.32	4.22	4.49
22	Environmental impact of EPIM	4.90	4.30	4.63
	Total	100.00	100.00	100.00

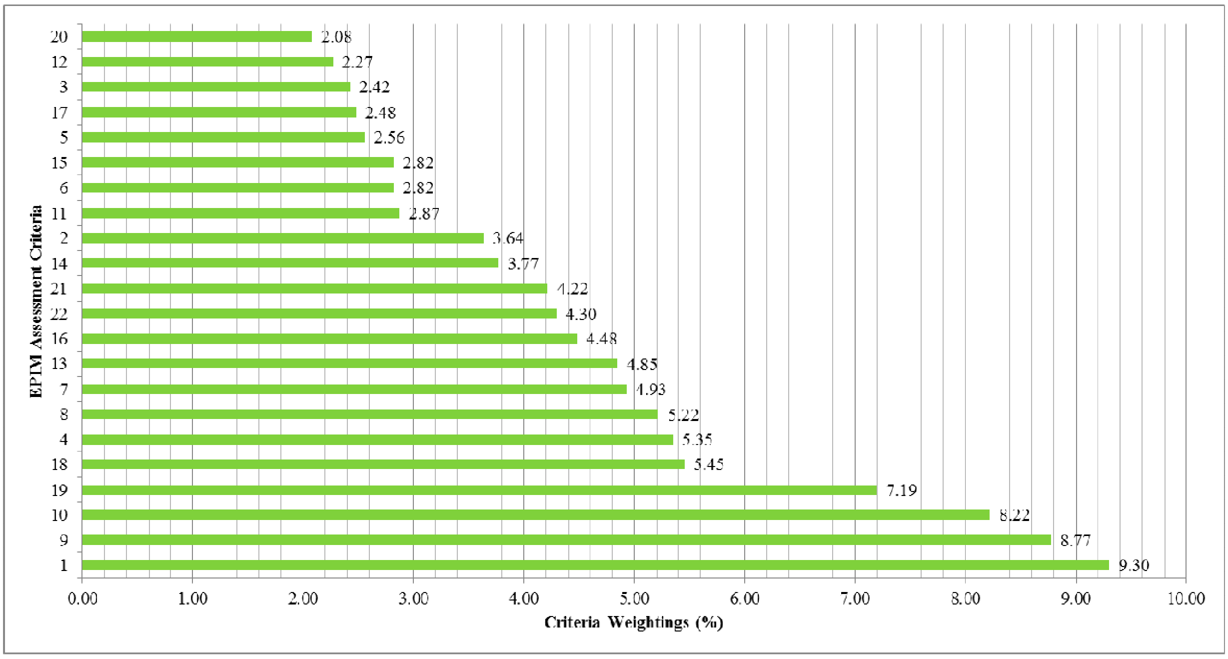

Figure 1 Criterion weightings obtained from bipolar sum of differences analysis

The prominence given to criterion 19 'impact on building's vapour permeability' is unexpected because this is a very technical attribute of an EPIM. It can be explained because it is a key feature of the traditional form of construction that the survey has targeted. Traditional buildings are known to be more complex in the way in which the building fabric manages moisture, as compared to modern impermeable construction forms [18]. Therefore, any EPIM that directly impacts the building fabric and how it interacts with the environment could have a detrimental effect upon the original building, as well as reducing the EPIM's performance, if it is not compatible with or not appropriately applied to the traditional construction form. One expert stated "...traditional buildings enable moisture to move through their fabric, get that wrong and you get an impervious building that gets damp and rotten quickly". Therefore criterion 19 essentially represents the EPIM's suitability for the construction form undergoing improvement. It would be interesting to see how highly this criterion would have been weighted if the assessment criteria were designed to address an existing building of modern construction.

The contrast between the importance placed on the fabric and the services is quite marked. Criterion 17 'impact on existing building services' is weighted lowly (position 19 in the rank order) whereas criteria 19 and 4 clearly relate to the effect of an EPIM on the building fabric, and are ranked in the top 6. This could again be due to the focus of this survey on traditional construction, and the more complex behaviour of the fabric than in modern construction.

Of the criteria with low weightings, numbers 3 and 12 concern the ease of installation and maintenance of the EPIM, respectively. The low importance is probably because these works would most likely be carried out by external contractors and therefore the risks are passed on to them. Cost of disposal of the EPIM at the end of its life (criterion 20) is considered relevant but of low importance, perhaps because this is a cost to be borne in the future. Disruption to the occupants during installation (criterion 6) can be managed by working out of hours, and training the occupants in the use of the new systems (criterion 15) is possibly less important when handing over a refurbished traditional building than when first occupying a new low-energy building.

Space precludes a detailed examination of the different perspectives of the experts, but it is clear that, while all the experts weighted capital cost, financial payback, potential energy savings and impact on the vapour permeability of the building fabric highly, those with a heritage focus also weighted loss of significant fabric and impact on appearance highly. Reliability of the EPIM was weighted the same but less highly by all experts. Experts with a client focus weighted internal comfort, existing services and impact on internal air movement more highly than other expert groups.

Conclusions

A crucial step in energy-led refurbishment of buildings is the multiple criteria decision-making process that must be applied to individual energy performance improvement measures. A Delphi survey identified 22 criteria, covering the installation, operation and disposal stages of the life-cycle against which such decisions are made. Pairwise comparisons using an group of experts enable weightings of the relative importance to be assigned to each criterion, and this can be used by decision-makers to score the refurbishment options in terms of appropriateness to their own building.

The Delphi technique is suitable for this purpose. However, whilst an expert survey with 13 respondents is statistically acceptable, a larger group would have been preferable and should be targeted in any future work. Further work should investigate how the relative weightings vary between expert groups with different perspectives and should extend the study to residential refurbishment.

Acknowledgements

We are grateful to the 13 experts which gave time to participate in the Delphi survey. The work was financed by the Engineering and Physical Sciences Research Council under a Doctoral Training Award.

References

[1] Strachan M.E., Banfill P.F.G. *Decision support tools in energy-led, non-domestic building refurbishment: Towards a generic model for property professionals*. Facilities 30 (9), 2012, pp374-395.

[2] Yoon K.P., Hwang C.L. *Multiple attribute decision making: an introduction*. Sage University Paper series on Quantitative applications in the social sciences, 07-104. Thousand Oaks, CA: Sage Publications Inc 1995.

[3] Hon C.K., Chan A.P., Chan D.W. *Strategies for improving safety performance of repair, maintenance, minor alteration and addition (RMAA) works*. Facilities, 29(13/14), 2011, pp591 - 610.

[4] Eskandari H., Sala-Diakanda S., Furterer S., Rabelo L., Crumpton-Young L., Williams K. *Enhancing the undergraduate industrial engineering curriculum: Defining desired characteristics and emerging topics.* Education+Training, 49(1), 2007, pp45 - 55.

[5] Bond S., Bond J. *A Delphi survey of clinical nursing research priorities*, Journal of Advanced Nursing, 7(6), 1982, pp565 - 575.

[6] Schmidt R., Lyytinen K., Keil M., Cule P. *Identifying software project risks: An international Delphi study.* Journal of management information systems, 17(4), 2001, pp5 - 36.

[7] Lunsford D.A., Fussell B.C. *Marketing business services in central Europe: the challenge: a report of expert opinion*. Journal of Services Marketing, 7(1), 1993, pp13 - 21.

[8] Höjer, M. *Transport telematics in urban systems—a backcasting Delphi study.* Transportation Research Part D: Transport and Environment, 3(6), 1998, pp445 - 463.

[9] Quade E.S. *On the Limitations of Quantitative Analysis* [online]. Available from: http://www.rand.org/pubs/papers/P4530.html [accessed 14 January 2014].

[10] McCauley C. *The nature of social influence in groupthink: Compliance and internalization.* Journal of Personality and Social Psychology, 57(2), 1989, pp250-260.

[11] Hsu C.C., Sandford B.A. *The Delphi technique: making sense of consensus*. Practical Assessment, Research & Evaluation, 12(10), 2007, pp1 - 8.

[12] Oh K.H. *Forecasting through Hierarchical Delphi*. PhD. thesis, 1974. Ohio: The Ohio State University.

[13] Pill J. *The Delphi Method: Substance, Context, a Critique and an Annotated Bibliography.* Socio-Economical Planning Sciences, 5, 2007, pp57 – 71.

[14] Turoff M. *The design of a policy Delphi.* Technological Forecasting and Social Change, 2, 1970, pp149 – 171.

[15] Bartlett C.J., Heermann E., Rettig S. *A Comparison of Six Different Scaling Techniques.* The Journal of Social Psychology. 51, 1960, pp343 – 348.

[16] Hajkowicz S.A., McDonald G.T., Smith P.N. *An Evaluation of Multiple Objective Decision Support Weighting Techniques in Natural Resource Management*. Journal of Environmental Planning and Management, 43, 2000, pp505 – 518.

[17] Hamilton L. *Creating environmental weightings for construction products – Results of a study*. IHS BRE Press, 2007. ISBN 978-186081968-1.

[18] CIBSE. *Guide to building services for historic buildings*. Chartered Institute of Building Services Engineering. London, 2002. ISBN 1-903287-30-8.