An investigation into Building Information Modelling Assessment Methods (BIM-AMs)

Ammar Azzouz, MSc. Alexander Copping, Ph.D. and Paul Shepherd, Ph.D.
University of Bath
Bath, UK

This paper reports on an investigation into the mapping of Building Information Modelling Assessment Methods (BIM-AMs) in the Architecture, Engineering and Construction (AEC) sector. Drawing on insights from the literature, twelve BIM-AMs were initially analysed and compared. Some of the key points distinguishing AMs are the level of detail and areas of measurement. Based on the fact that BIM-AMs span the full range of the complexity spectrum, two AMs were used in a number of case study projects to explore their variations. The first, a simplified AM, is the National BIM Standard Capability Maturity Model (NBIMS-CMM) developed by the National Institute of Building Sciences (NIBS). The second, more detailed, is the Virtual Design and Construction (VDC) Scorecard, developed at Stanford University. A comparative case study methodology was implemented to investigate characterisations of the two AMs. The analysis included two different phases. The first phase assessed the reliability of using one BIM-AM on a single project but was completed by two members of the same team to explore subjectivity. The second phase applied two BIM-AMs to the same project to investigate whether they gave consistent assessments. Observations of the results identified limitations of BIM-AMs, in particular in their dependency on qualitative judgements. They are beneficial for industry, stakeholders and decision makers to measure BIM implementation, but attention needs to be given to the possibility of using more quantitative measurements in assessing BIM maturity.

Keywords: Assessment Methods, BIM, NBIMS-CMM, VDC Scorecard

Introduction

Implementation of Building Information Modelling (BIM) poses a fundamental challenge for the AEC industry as ‘an environment of change’ (McCuen, Suermann, & Krogulecki, 2012, p. 224). In the UK, unlike many other countries, Government strategy will require ‘collaborative 3D BIM’ for publicly funded projects from 2016. According to the ‘NBS National BIM Report 2014’, a survey of more than 1,000 construction professionals in late 2013 indicated that only half were using BIM. Findings showed that over 95% of participants were aware of BIM, but only 54% of those aware of BIM were actually using it (RIBA Enterprises, 2014). Whilst a significant amount of research has been carried out on the implementation of BIM in the AEC industry, far less attention has been paid to the evaluation of BIM utilisation. This paper contributes to that modest, but increasing, literature by investigating twelve existing BIM Assessment Methods (BIM-AMs). Most literature on this research topic focuses on presenting current AMs independently. In contrast, this paper provides a collective comparison of these AMs to explore their theoretical and methodological design.

For more than a decade, a considerable body of research has been directed to support and guide the industry in its implementation of BIM. However, with the emergence of BIM, few frameworks have been developed to assess BIM’s level of maturity (Change Agents AEC, 2013; CIC, 2013). Among the most recognisable of these AMs is the National BIM Standard Capability Maturity Model (NBIMS-CMM), at least in terms of being the first to be developed. The tool was launched in 2007 to evaluate BIM adoption in projects (NIBS, 2007). Since then, a number of other AMs have been developed to support and assess BIM applications (Sebastian & Berlo, 2010). This paper analyses briefly the similarities and differences between the available BIM-AMs. It presents further understanding of the performance of these AMs in practice by applying two of them to a number of projects. The emphasis, therefore, is to evaluate AMs that evaluate projects and not the projects themselves. This differs from existing research in two ways. Firstly, instead of concentrating on presenting each of the available BIM-AMs separately, the investigation compares and explores the variations between these AMs. In contrast to previous studies, the concern of this paper is to outline the relationship between these AMs by providing a visual representation of their progression. Secondly, rather than taking one BIM-AM to evaluate a project, a comparative case study method was used to explore how similar or different the performance of BIM-AMs is in two phases including; phase 1: two different members applying one AM to a same project and phase 2: applying two AMs to the same project.

Current BIM Assessment Methods (BIM-AMs)

Research in the area of BIM-AMs has matured noticeably since BIM was first coined as a term more than a decade ago. Over this period, the BIM field has witnessed a growing body of development in the number of methods and tools developed by academics and professionals (CPI, 2011; VICO, 2011). Since the first AM was developed in 2007, research into BIM-AMs has evolved increasingly, but is still lagging behind other AMs, such as green building assessments (Kam, Senaratna, Xiao, & McKinney, 2013b). Some of these BIM-AMs have more significant strengths and weaknesses than others; for example, some tools are more user-friendly, easy to apply and practical, whilst others are difficult to use and require more time to complete. The literature of BIM-AMs focuses on either software or procedural aspects of the BIM implementation process (Succar, 2010). NBIMS-CMM, which was developed in 2007, focused only on the technical part of the BIM process. However, while the development of AMs has proliferated gradually since then, more recent research has observed the limitations of the previous tools and addressed more factors to be measured. Consequently, in the past few years, BIM-AMs have shifted from evaluating one aspect of BIM to assessing wider measures related to people, business processes, technology, communication and information exchange. This paper attempts to address the limitations and challenges of these AMs by presenting a brief history of the development in the BIM-AMs field.

From the perspective of this paper, current BIM-AMs have a number of limitations. Three of them are posed; inaccessibility, absence of documentation and subjectivity. Firstly, some of the current BIM-AMs are not readily accessible in terms of methods development, because of public access and cost issues. bimSCORE (2013) is an AM that provides a brief free assessment online (FREE 004 version). However, it requires a fee to access its ‘NOW 010’ version, which provides written and concise results on BIM maturity. Another issue related to inaccessibility is the restriction on some of the AMs to be used only within their original research team, due to their commercial value. The Owner’s BIMCAT is the latest AM, developed by Giel (2013) for her PhD at the University of Florida. Unlike other researchers on BIM-AMs, Giel provided a comparison of the existing AMs presenting areas of similarities and differences. However, the thesis itself is inaccessible until 2015. The second limitation of the current BIM-AMs is the lack of documentation where some of the methods were developed and no further research was carried out. The ‘CPIx BIM Assessment Form’, the only AM developed in the UK, is available on-line in PDF format, but it lacks publications describing the concept of the tool and the certifying process (CPI, 2011). Similarly, another AM, was developed a number of years earlier, ‘BIM Proficiency Matrix’, and was not supported with further documentation (Indiana University, 2009). However, the main limitation is the absence of publications on the conceptual and methodological approach of this tool. In general, the reasons behind these two barriers, the inaccessibility and lack of publication, could be due to agreements that some researchers sign with their funders, or the lack of support and funding for further studies on their work.

The third limitation of the existing BIM-AMs is the inherent problem of subjectivity. Research on BIM-AMs relies on high levels of qualitative measurements to evaluate BIM implementation. Factors of measurement included in these tools are subjective and qualitative in their manner. Exceptions can be found in only a few of the identified AMs where researchers attempted to include both qualitative and quantitative measurements (VDC Scorecard and Characterisation Framework). They addressed quantitative factors such as time spent on creating model and number of stakeholders creating and reviewing it. A major criticism of subjective measures is that they rely on the personal judgement of participants, which is influenced by their level of knowledge and current status in the organisation, whereas objective measures are more quantitative and exact (Nicoletti & Pryor, 2006).

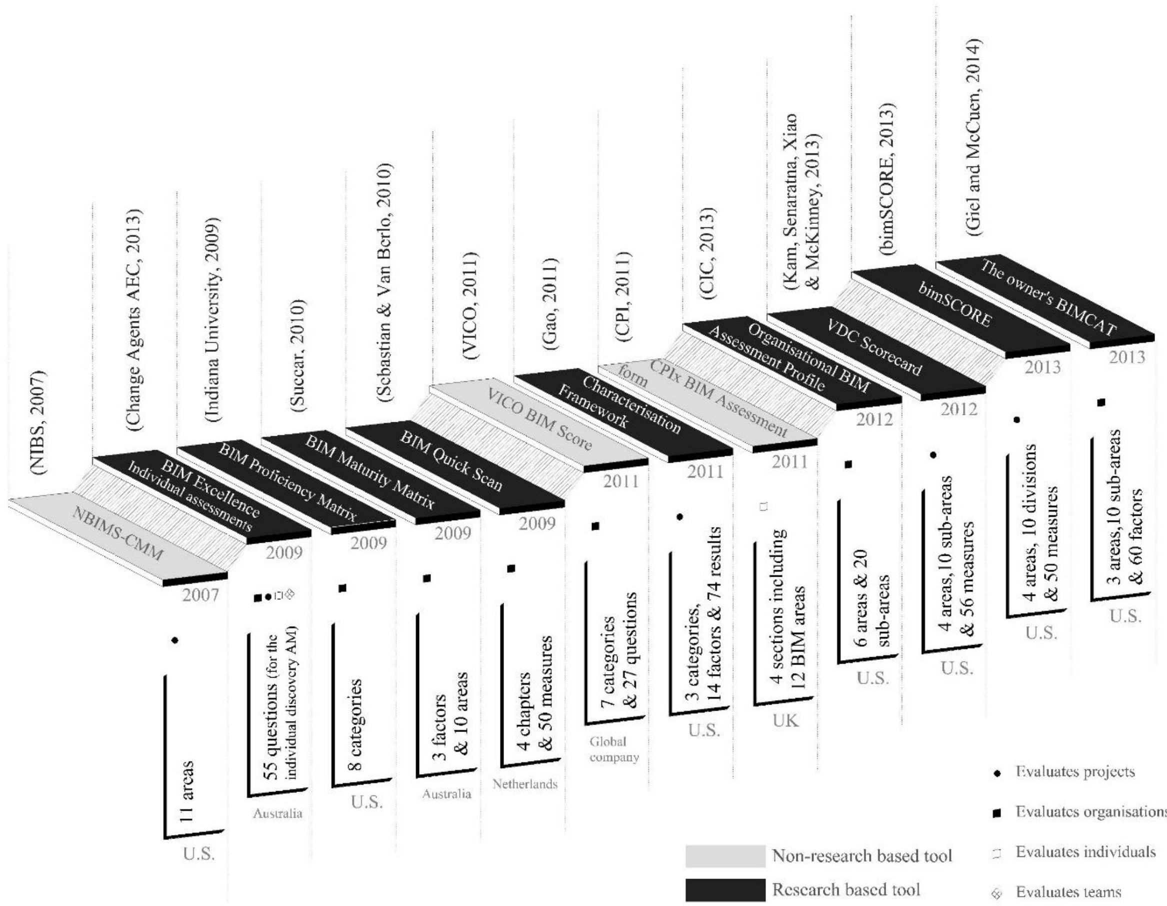

Figure 9: Existing BIM-AMs

The authors of this paper illustrated the brief history of BIM-AMs in Figure 1. At least twelve such AMs were developed between 2007 and 2014. The emergence of these AMs reflected a growing research on this concept as a vehicle for classification in the BIM domain. Many differing properties are presented including the country of origin, whether they are a research or non-research based tool, the number of BIM measures, and whether the AM evaluates a project, team, individual or organisation. One of these differences observes the 'complexity versus simplicity' provided by these tools. Developers of BIM-AMs identified different levels of detail depending on the number of factors and information needed for each assessment. Half of the current tools identify more than 50 measurements for their methods (Figure 1). For instance, the BIM Characterisation Framework, developed by Gao (2011) at Stanford University, includes 3 categories, 14 factors and 74 results. What is of concern with such detailed methods is the time needed to complete the assessment and, more importantly, the ability to provide accurate answers to a wide range of questions. When applying the VDC scorecard to 108 pilot projects, the average time for the interviewee to complete the Express Version of this AM was 4 hours (Kam, Senaratna, Xiao, & McKinney, 2013a), whilst the average number of answered questions was 72% (Kam et al., 2013b). In contrast, the other half of the AMs was concerned with shorter forms of evaluation and required less information. The Organisational BIM Assessment Profile AM, developed at Pennsylvania State University (CIC, 2013), provides 6 areas and 20 sub-areas as factors to measure. However, although these tools are short and included definitions of their terms, some of them such as NIBIMS-CMM, were criticised for not being easily understood (Succar, 2010). The link between complexity and simplicity is found by the authors of this paper to be dynamic. Thereby, simplicity might be found in 'complex' methods. This claim could be proved when a detailed method employs clear language, description of the measurements and a structured framework. It also found that demystified and short tools might be difficult to apply.

Methods

Two of the current AMs were applied in industry to explore their performance, the National BIM Standard Capability Maturity Model NBIMS-CMM Version 2 (NIBS, 2012), and the VDC Scorecard Version 1.3 (CIFE, 2013), presenting varied characteristics. The reason behind selecting these AMs is that the emphasis of this paper is to evaluate 'projects', and not 'organisations', 'teams' and 'individuals'. The choice of evaluating 'projects' in this report lies in the need to provide a snapshot of BIM implementation in the organisation level and more broadly in the industry level. This limited the options to the five project based tools out of the 12 existing AMs (Figure 1). A major criticism of these five AMs is the absence of documentation on how to apply them. Therefore, the two chosen AMs are currently the only free and well documented tools available to evaluate BIM level of maturity of 'projects'. There are a few similarities between the two selected tools. Both developed in the U.S., evaluate projects and provide a five level scoring system. However, they differ in a number of important respects. NBIMS-CMM is a simplified, brief, non-research based tool. Conversely the VCD Scorecard is a detailed, comprehensive and a research-based AM.

Phase 1: Two Different People Applying One BIM-AM to a Single Project

A number of authors signal the use of NBIMS-CMM in practice (McCuen et al., 2012; Sebastian & Berlo, 2010), but no study has applied this AM to the same project independently by two different members of the design team. After developing the tool in 2007, the NIBS had to ensure that it could be used successfully. Therefore, the NBIMS Team evaluated nine award winning projects. Each single project was tested by six assessors and the differences between the scores were no more than 5% (NIBS, 2007). However, in this paper, the authors applied a different testing method. Instead of evaluating the project by a number of independent assessors, the same project was assessed by two participants who had worked on the project being evaluated. The reason behind this lies in the need to observe how subjective or objective the AM is in practice.

Assessment of an 'Opera House' project was carried out in association with an award-winning structural engineering practice based in the UK. In this case a senior structural technician (Participant 1), and a graduate engineer, (Participant 2), were both using BIM on the Opera House, which was at the time of the interview (27/07/2014), moving into the construction stage. Both participants were unaware of the existence of BIM-AMs before being approached by the authors. Therefore, a brief background of BIM-AMs was introduced prior to the assessment. BIM technology was implemented in this project as a unified platform between all stakeholders (using Autodesk Revit). Each of the participants took nearly 20 minutes to complete the hard copy of the NBIMS-CMM assessments. Their answers were later applied to the Interactive CMM (I-CMM) by the authors. A table and a diagram, (Figure 2), were automatically generated, providing visual representation of the scores. The overall score that Participant 1 gave (52.8%) differed significantly from Participant 2 (41.8%).

Physically, the two participants were located in the same office, sitting next to each other and working on the same project. Despite this close working relationship, their scores were noticeably different when they completed the same assessment of the same project. Table 1 compares the score distribution of the 11 measured areas given by participants, and highlights, in grey, categories of equal rank order. The overall score given by Participant 1 is nearly 10 points higher than Participant 2, reflecting how different results can be given on the same project. Answers for almost all the categories provided by the two participants were not similar, except in two cases; 'Business Process' and 'Interoperability'. These areas were scored by both participants as 2.7 and 5.8 respectively. The most remarkable difference between the scores could be found in the 'Roles and Disciplines' category, with a discrepancy of nearly 3 points. Despite the descriptions defining each of the 11 areas of interest, the two participants understood some of the terms differently. Participant 2 argued that areas of interest in this AM were subjective and not numeric as in BREEAM. Therefore, it was very likely that both answered differently.

Table 7

Capability Maturity Model Scores for the same project with different participants

Area of Interest	Rank order	**Participant 1**	Rank order	**Participant 2**
Data Richness	3)	5.9	1)	6.7
Life-cycle Views	8)	3.4	8)	2.5
Change Management	7)	3.6	10)	1.8
Roles or Disciplines	5)	5.4	7)	2.7
Business Process	10)	2.7	7)	2.7
Timeliness/ Response	6)	4.6	5)	3.6
Delivery Method	6)	4.6	6)	2.8
Graphical Information	1)	7.4	2)	6.5
Spatial Capability	9)	2.8	9)	1.9
Information Accuracy	2)	6.7	4)	4.8
Interoperability/ IFC Support	4)	5.8	3)	5.8
Overall Score		52.8		41.8

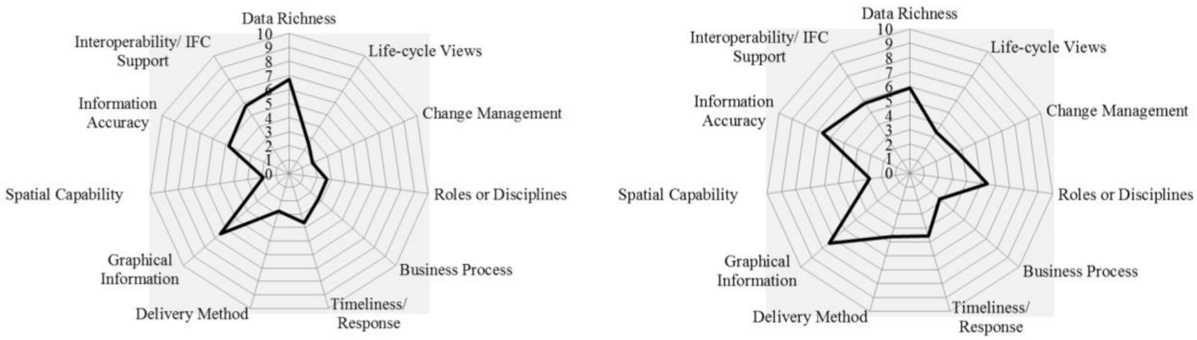

Figure 2: Areas of Interest Diagram - Completed by Participant 1 (left), and Participant 2 (right) (Adapted from: NIBS, 2007)

In comparing the data provided by both participants, it seems that both share roughly similar expectations of areas of strengths and weaknesses. This could be seen when comparing the two 'areas of interest charts' in Figure 2. The resulting scores seem to follow a similar pattern, with both participants rating the same criteria high and low, albeit with different absolute score values. One of the highest scores that both participants gave was in the 'Graphical Information' category (scoring 6.5 and 7.4 respectively). This means that both agreed that they were providing all stakeholders with '3D object-based' models. Another example of similar ranking order can be found in the 'Life-cycle views', 'Delivery Method' and 'Spatial Capability' area, where both of them scored it as the 9th weakest category in their answers (highlighted in grey in Table 1). What is of concern is that the same project achieved two different absolute scores from two participants working in the same office. This raises the question of how valid this AM is. The 'Opera House' case draws attention to the need for an accurate BIM-AM where projects can be evaluated objectively by different employees and still lead to a similar score.

Phase 2: Applying Two BIM-AMs to the same Project

This section describes the implementation of two different BIM-AMs on the same project. Most studies of the use of BIM-AMs focused only on implementing one method, but no single study exists which adequately assesses the same project by two different BIM-AMs. Therefore, both the NBIMS-CMM and the VDC Scorecard were applied to the same project in order to examine whether they gave consistent assessments. This scenario was implemented in two case study projects. It is important to note that the VDC Scorecard, (Version 1.3), is a very detailed tool, which includes four main areas: planning, adoption, technology and performance. To complete the full assessment a couple of hours were required. However, for this paper only the 'technology' part of the tool was completed, since it was difficult to find participants who were willing to spend two hours to complete an assessment. Additionally, the selected part of the AM is the closer to the areas assessed by NBIMS (Table 4 presents areas of similarities between the two AMs).

The first case study project in this phase was a 'private schools' project in a global consulting organisation that provides planning, developing, design and construction services in the UK. A principal BIM integrator completed the two different AMs assessments for the same project. The project consists of a group of schools and, at the time of the interview, (04.06.2014), it was moving to the construction phase. However, the answers provided by the participant were based on the schematic design phase. The Software application used for this project was mainly Revit, and drawings were saved as PDFs to share them with other stakeholders. When completing the NBIMS-CMM, the overall score was 52.9, points obtaining a 'Minimum BIM' as a level of maturity, with 7.2 more points required to achieve BIM certified (see Table 2). Initial analysis of the data provided by the participant for this project scored high levels for 'Graphical Information', 'Change Management' and 'Information Accuracy'. Far fewer points were allocated to the other categories.

In a second stage, the same project was evaluated again by the same participant using another AM, the VDC Scorecard, in the 'technology' area. The overall score in this case was 66%, which brought it to 'Advanced Practice' classification as seen in Table 2. Areas of measurements in this AM were varied, detailed and in a few occasions numeric. An example of these questions is the percentage of 'information loss after model exchange'. For this specific answer, the participant pointed out confidently that there was 0% loss of information while sharing models between architects and engineers. An explanation for this was the fact that all members involved in this project used the same piece of software (Autodesk Revit) covering differing uses such as 3D modelling and rendering. From this model, DWG and PDF files were extracted and then circulated at project management level providing construction and cost documents.

Table 8

Levels of Scoring System in NBIMS-CMM and VDC Scorecard

NBIMS-CMM (Version 2)	Minimum BIM (40-59.9)	Certified (60-69.9)	Silver (70-79.9)	Gold (80-89.9)	Platinum (90-100)
VDC Scorecard (Version 1.3)	Conventional Practice (0-25%)	Typical Practice (25-50%)	Advanced Practice (50-70%)	Best Practice (70-85%)	Innovative Practice (85-100%)

A similar scenario was seen in another case study project, a university building project in another practice located in the UK. Results in this case were roughly similar to the findings in the previous case study and, therefore, are briefly discussed here. The project was assessed applying both NBIMS-CMM and VDC Scorecard. In the first stage, the project scored an overall of 51 points when applying NBIMS-CMM, which categorised it to 'Minimum BIM'. This AM was completed by Participant A, who ran the project. However, Participant B, an architect, completed the VDC Scorecard on the same project allocated an overall score of 64% and classified it as 'Advanced Practice' (Table 2). CIFE researchers established a 'confidence level' in case participants were uncertain about their answers. This was applied in this case with a confidence level of 83%. The findings of the two case studies can be seen in Table 3.

The two cases above provide examples of applying two BIM-AMs on the same project. In both cases the scores achieved by the two methods were significantly different (Table 3). When NBIMS-CMM was applied to either project, a 'Minimum BIM' level was achieved, which is the lowest level the tool provides. In contrast, when assessing the same project with the VDC Scorecard, the project ranked 'Advanced Practice' in both cases, which is the third level of maturity out of five (Table 2 highlighted in grey).

Table 3

Comparing the Results of Two AMs on Same Project

BIM-AM/ Case study	Private Schools Project	University Building Project
NBIMS-BIM	52.9 points, 'Minimum BIM'	51.0 points, 'Minimum BIM'
VDC Scorecard	66%, 'Advanced Practice'	64%, 'Advanced Practice'

This rather contradictory result may be due to a number of differences. Two main causes of this result could be suggested; the first is related to issues the simplicity and complexity of BIM-AMs. NBIMS-CMM is a short, simplified tool whereas the VDC Scorecard is much more detailed and requires more time and information to be completed. In NBIMS-CMM, 11 areas were assessed as seen in Table 4. For each of these areas, participants had to choose one maturity level ranging between 1 to 10, (10 is the highest maturity). Thereby, participants had to provide only 11 non-numeric answers to complete the AM. In contrast, the VDC Scorecard, 'technology' section, assesses 12 areas, although five of them are actually included in the calculation of the overall score, namely highest level of detail, data sharing method, information loss after model exchange, starting phase and ending phase. Unlike the NBIMS-CMM, each area covers 19 factors. This means that respondents were required to provide 19 × 12 = 228 answers for VDC Scorecard compared to 11 in NBIMS-CMM. The second possible cause of the different scores achieved by the two AMs is the difference in categories assessed. The two BIM-AMs evaluated different areas of technology. 'Information Loss After Model Exchange', which is one of the main principles of BIM, was assessed in VDC Scorecard but not in the NBIMS-CMM. Overlap between the 'areas of measurement' of the two AMs is identified only in three categories (highlighted in Table 4).

Table 4

Comparison of Areas of Interest between VDC Scorecard and NBIMS-CMM

NBIMS-CMM	**VDC Scorecard**
Data Richness	Highest Level of Detail
Delivery Method	Data Sharing Method
Interoperability/ IFC Support	Model Exchange Format
Life-cycle Views	Model Uses Utilised
Change Management	Primary Software
Roles or Disciplines	Secondary Software
Business Process	Information Loss After Model Exchange
Timeliness/ Response	Starting Phase
Graphical Information	Ending Phase
Spatial Capability	Number of Stakeholders creating or Using Model/ File
Information Accuracy	Number of Members creating or Using Model/ File
	Stakeholder Leading Effort

Conclusions

While the adoption of BIM has rapidly developed as a 'hot topic' in research and industry, it remains for many vague both as a term and in its practical application. This absence of clarity lies in the lack of a framework to accurately define the boundaries of the term and assess its implementation. As a response, many scholars have developed methods to benchmark BIM capacities. This study contributes to this growing, yet small, body of research by comparing the current BIM-AMs. This analysis suggests a new starting point in this field, 'BIM-AMs' as an acronym, referring to BIM Assessment Methods. The main limitation of the contemporary BIM-AMs is that they are primarily concerned with qualitative data. Almost all the available AMs establish questions to which the answers cannot be given in clear and quantifiable terms. Respondents are not asked to provide numbers or measurements except in two AMs (VDC Scorecard and Characterisation Framework). This has been demonstrated when one of the tools was applied in a case study project with two participants. It is likely that no two members of the same team and working on the same project will necessarily have the same answers to the different questions in the AM. This is because definitions used in NBIMS-CMM, for instance, are 'subjective and open to interpretation' (NIBS, 2007, p. 76). However, even with the numerical approach of the VDC Scorecard, some of the questions are difficult to answer, as people have to 'guess' how many hours they spent on their models, if they have no precise tracking system. Therefore, the VDC Scorecard developers created the 'confidence level' as evidence. However, in order to achieve accurate measurements, an efficient evaluation system should capture quantifiable and objective metrics. The authors observed that there is a dependency on qualitative judgements in most BIM-AMs. Therefore, more attention needs to be drawn to the application of quantitative and objective measurements in order to achieve more efficient BIM-AMs. Whether used to assess projects or not, AMs will ultimately provide a further understanding of how the built environment is being designed, managed and constructed.

Acknowledgement

Acknowledgement is given to Feilden Clegg Bradley Studios, Integral Engineering Design and Parsons Brinckerhoff who contributed their time and experience to the case study projects.

References

bimSCORE. (2013). bimSCORE [WWW document]. URL https://www.bimscore.com/index.php/welcome

Change Agents AEC. (2013). BIM Excellence Performance Measurement Coporate Services

CIC. (2013). BIM Planning Guide for Facility Owners. Version 2.0. [WWW document].URL http://bim.psu.edu/

CIFE. (2013). CIFE VDC Scorecard Research Website [WWW document]. URL http://cife.pbworks.com/w/browse/#view=ViewFolder&param=Dimi

CPI. (2011). CPIx-BIM Assessment Form [WWW document]. URL http://www.cpic.org.uk/wp-content/uploads/2013/06/cpix_-_bim_assessment_form_ver_1.0.pdf

Gao, J. (2011). *A Characterization Framework to Document and Compare BIM Implementations on Construction Projects.* Stanford University.

Giel, B. (2013). A Comparison of Existing BIM Assessment Tools [WWW document]. URL projects.buildingsmartalliance.org/files/?artifact_id=5140

Indiana University. (2009). Building Information Modeling (BIM) Guidelines and Standards for Architects, Engineers, and Contractors. Kam, C., Senaratna, D., Xiao, Y., & McKinney, B. (2013a). The VDC Scorecard: Evaluation of AEC Projects and Industry Trends.

Kam, C., Senaratna, D., Xiao, Y., & McKinney, B. (2013b). The VDC Scorecard: Formulation and Validation. *Stanford University*.

McCuen, T. L., Suermann, P. C., & Krogulecki, M. J. (2012). Evaluating Award-Winning BIM Projects Using the National Building Information Model Standard Capability Maturity Model. *Journal of Management in Engineering, 28*(2), 224-230. doi: 10.1061/(ASCE)ME.1943-5479.0000062

NIBS. (2007). National Building Information Modeling Standard: Version 1.0- Part 1: Overview, Principles, and Methodologies.

NIBS. (2012). National BIM Standard- United States Version 2.

Nicoletti, G., & Pryor, F. L. (2006). Subjective and objective measures of governmental regulations in OECD nations. *Journal of Economic Behavior & Organization, 59*(3), 433-449.

RIBA Enterprises. (2014). NBS National BIM Report 2014 [WWW document]. URL http://www.thenbs.com/pdfs/NBS-National-BIM-Report-2014.pdf

Sebastian, R., & Berlo, L. (2010). Tool for benchmarking BIM performance of design, engineering and construction firms in the Netherlands. *Architectural Engineering and Design Management, 6*(4), 254-263.

Succar, B. (2010). Building information modelling maturity matrix. *Handbook of research on building information modelling and construction informatics: Concepts and technologies, J. Underwood and U. Isikdag, eds., IGI Publishing*, 65-103.

VICO. (2011). Calculation your BIM score [WWW document]. URL http://www.vicosoftware.com/resources/calculating-bim-score/tabid/273811/Default.aspx