The Maturity of Usability Maturity Models

Carmen L. Carvajal[(✉)] and Ana M. Moreno

Facultad de Informática, Universidad Politécnica de Madrid,
Campus de Montegancedo S/N Boadilla Del Monte, Madrid, Spain
carmen.carvajal.jimenez@alumnos.upm.es,
ammoreno@fi.upm.es

Abstract. The integration of usability practices in software development is not a straightforward process. In this context, the application of usability maturity models (UMM) in a software organization can provide insightful information to improve such integration. This paper discusses the design and application characteristics of the UMMs used over the last decade. The analysis of recent UMMs confirms that, even when the UMM field is a matter of interest and is getting adapted to new development contexts (for instance, agile or open source). UMMs lack detailed empirical evidence and supporting documentation for their objective application. In addition, our study also identifies other open issues related such as the level of prescriptiveness or mutability of UMMs. Consequently this paper identifies different opportunities for improving the maturity of UMMs. The application of mature UMMs would contribute to a better incorporation of usability and user experience practices in software organizations.

Keywords: Usability maturity models · UX maturity models · Maturity models

1 Introduction

Usability is a quality attribute defined in ISO/IEC 25010 [1] as the degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. The benefits of adopting usability or UX (user experience) practices in software development have been much highlighted from the viewpoint of both users and organizations [2, 3]. However, usability methods are hardly ever integrated into software lifecycles in industrial settings [4]. In this context, usability or UX maturity models (UMMs, from now on) have an important role to play in such integration.

According to Becker et al. [5], a maturity model provides the criteria and characteristics that need to be fulfilled in order to achieve a particular maturity level in a specific area. The aim of maturity models in the field of usability is to evaluate the maturity of an organization from a usability point of view. That is, UMMs help to reflect how the usability process and practices are implemented in an organization. As a result of the evaluation, the organization can identify which aspects of usability require improvement. UMMs are thus a very useful tool for an organization to improve its software process from a usability point of view.

There are several UMMs in the literature. Conversely, there is not much research analysing their characteristics and practical applicability. In 2012, Wendler [6] published a systematic mapping study of maturity models applied to different domains. However, he refers only briefly to the specific case of UMMs as part of his discussion of the application of maturity models in the field of software. As far as we know, the only study on UMMs was published in 2006 by Jokela et al. [S1] Jokela et al.'s survey identified 11 pre-2006 UMMs.

A decade later, our aim is to gain an up-to-date snapshot of the state of UMMs in order to identify valid models and their characteristics from both the structural and application viewpoints. To do this, we conduct a systematic mapping study of the UMM literature published over the last ten years; this includes publications about UMMs object of study during the last decade even if they have been originally published before. Practitioners looking to improve the adoption of usability in their development process may find the results useful, as they paint a picture of current UMMs together with their potential strengths and weaknesses. This information is also useful for researchers, as it suggests open lines of research.

The results of this research are reported as follows. Section 2 discusses the research previous to ours. Section 3 explains the applied systematic mapping research method (including the research questions, search process and information extraction process). Section 4 details the results. Finally, Sect. 5 discusses the results and conclusions of this research.

2   Related Work

Wendler [6] published the first ever systematic mapping study in the literature on maturity models in 2012. The study revealed a growing interest in the topic with 237 articles retrieved from 1993 to 2010. This study identified 22 domains in which maturity models have been applied, including knowledge management, information management or IT governance. According to Wendler, software development and software engineering models are the leaders, as there are significantly more articles in this than in other areas. Due to the breadth of the study, however, it mentions UMMs only briefly as an example of one of the subdomains where models are applied in the software field. Wendler highlights weaknesses with respect to model validation and stresses the need to examine the suitability of maturity models in real scenarios.

In the specific area of UMMs, in 2006, Jokela et al. [S1] identified 11 models published up to 2005. Jokela et al. studied these models from the general and practical viewpoints, analysing, for example, the number of levels that they contain or the usability elements that they evaluate. Jokela et al. [S1] concluded that hardly any of the analysed UMMs provide specific guidelines for their practical application or have been empirically validated. Additionally, they suggested a need for a cumulative research tradition on UMMs to help identify problems with existing models, understand the differences between the different models and avoid redundancies.

More recently, Salah et al.[S11] compared 11 MMUs in 2014 aiming to select the appropriated model in order to evaluate the Usability Maturity level in organizations using Agile methodologies and User centered design. In that work, the authors

analyzed the models according to their applicability in agile environments. For instance, they required that the models should be light, that is, they do not require considerable amount of time in their evaluations so that they do not interrupt the dynamics of agile methodologies. As a result, although with limitations, the Corporate Usability Maturity, and the Usability Maturity Model-HCS were classified as the most suitable for their application in agile contexts.

3   Research Methodology

Our research follows the systematic mapping procedure designed by Peterson et al. [7]. The research methodology aims to provide an overview of a research area and identify the type and quantity of research conducted and the published results. This section introduces the followed steps for this study.

3.1   Identify Research Questions

As already mentioned, the main aim of this paper is to compile and analyse the studies published from 2006 to 2016 on UMM in order to gain an overview of the field. The research questions for this study are the following:

- RQ1. Which usability maturity models have been addressed by publications over the last 10 years?
- RQ2 What are the general features of the UMMs?
- RQ3 What are the design features of the UMMs?
- RQ4 What are the use features of the UMMs?

The first question (RQ1) identifies the UMMs that have been addressed by publications over the last decade. The other questions aim to delve deeper into these models. Firstly, RQ2 provides an overview of the models by analysing their general features, including their application domain or number of maturity levels used, according to the general analysis presented by Jokela et al. [S1]. RQ3 and RQ4 gather more detailed information on the models from the viewpoint of their structure and application features, respectively. For that aim, particular criteria defined for evaluating maturity models will be used.

3.2   Search Relevant Literature

Based on the above research questions, we defined a set of keywords for searching terms related to usability and usability maturity models. As a result, the search string used was: *(usability OR "human centred design" OR "user centred design" OR "user experience") AND ("usability capability maturity model" OR "usability maturity model" OR "usability maturity" OR UCMM).*

The search was conducted from June to September 2016. It was originally confined to the title and abstract fields of the papers. As the number of returned results was low, however, the search was finally extended to the entire paper. The following electronic databases were used: ScienceDirect, ACM digital library, IEEExplore and Springer

Link. We also queried the Scimago scientific journals ranking in order to make sure that we searched at least the top twenty journals listed under Q1 in the field of Human Computer Interaction – HCI. To do this, we used Google Scholar as a secondary search engine to retrieve information of interest from HCI journals like Cyberpsychology, Behavior, and Social Networking or Topics in Cognitive Science. Finally, we applied the backward snowballing sampling technique on the selected set of papers.

3.3   Select Relevant Papers Based on Inclusion/Exclusion Criteria

We screened the papers returned by the search based on the inclusion and exclusion criteria listed in Table 1. Table 2 summarizes the paper identification process. Initially, the database search using the defined searching-string returned 309 papers. After applying the basic exclusion criteria, the number of papers was reduced to 250. After screening by the title and abstract, 24 papers were left. At this stage of the process, the second author sampled the selected and excluded papers at random to confirm the results. Later, another three papers retrieved by means of the backward snowballing technique were added. The final decision on which papers were selected was taken after reading the full text of the paper. Finally, 17 papers that strictly met the objectives of our research were selected as primary studies (see Appendix A for the full list).

Table 1. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Papers published since 2006	Papers not written in English
Conference, workshop and journal papers published in the area of software development and HCI	Papers whose full text is not available
Papers describing UMMs	Redundant papers
Papers on the results of applying UMMs	Notes, electronic presentations, poster papers, comments, patents or letters
Papers comparing different UMMs	Papers not focusing on the area of UMMs

Table 2. Paper identification process

Phase	SD	Springer	ACM	IEEE	GS	Total
Search using search string	14	45	22	17	211	309
Exclusion based on basic criteria like language or redundancy	14	43	20	17	156	250
Exclusion based on title and abstract	1	12	2	2	7	24
Inclusion based on backward snowballing	1	12	2	2	10	27
Exclusion based on full text	1	8	0	1	7	17

3.4   Build the Classification Scheme

The articles are organized base on their title, authors, publisher and date. In addition, the articles are classified as primary [8] (i.e., contains the original information), or secondary [8] (i.e., contains information based on a collection of primary studies).

Finally, the MMU title and origin (i.e., Academia or industry) are identified from the article. In addition, the articles are classified as solution, validation, evaluation, experience, philosophy, or opinion according to the criteria presented in [9].

3.5   Extract Information and Map Studies

Figure 1 illustrates a breakdown of the papers according to the classification presented in Sect. 3.4. Nine (9) out of the seventeen (17) papers (S2, S3, S4, S5, S6, S7, S8, S9 and S10) focus on defining new UMMs and are classified as solution studies. They are followed by five evaluation papers (S5, S11, S13, S14 and S17) and another two opinion papers (S15 and S16). Only one validation paper (S5), one experience paper (S12), and one secondary study (S1) have been published. Note that one paper (S5) was classified in three different categories, as it reported evidence related to a solution, validation and evaluation (for this reason, the number of papers listed in Fig. 1 totals 19). Note that the research tool used in the evaluation papers was case studies.
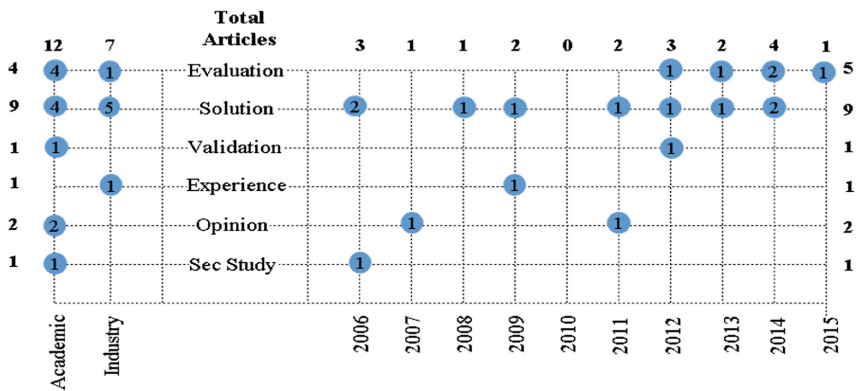

Fig. 1. Map by years and study setting

Additionally, Fig. 1 shows the distribution of the publications by year. This distribution was quite similar from year to year, except in 2010 when no papers were published and 2014 when slightly more papers were published. Figure 1 also shows the distribution of these publications by their setting: academia or industry. There are more papers from academia. This applies to all study types, except for experience papers, where we identified only one primary study conducted in industry, and solution studies, where the five studies from industry illustrated in Fig. 1 refer to four different models.

Although there are more papers from academia —the ratio is about 60 to 40—, we have found that interest in industry is more significant than in other software areas where a much smaller percentage of papers are sourced from industry like example [10].

Finally, the reader could think this study has some threats to validity. The first threat is that only four digital libraries were used (ACM, IEEE, ScienceDirect and Springer Link); however, according to Petersen et al. [17], these are the most relevant libraries for this subject and the use of IEEE and ACM as well as two indexing databases is sufficient for this research. Secondly, the reader could argue that not every study was taken into consideration for this work; however, several strategies were applied in order to mitigate this threat such as the keywords selection and the application of the backward snow-balling sampling technique on the selected set of papers. Lastly, in order to counteract any subjective bias on the part of the first author, the final decision on which papers were selected was taken after reading the full text of the paper. Additionally, all the papers were reviewed by the second author separately. The results of the evaluation were compared, and disagreements were settled by negotiation.

4   Research Results

This section presents answers for the research questions stated above.

4.1   RQ1. Maturity Models Under Research Over the Last 10 Years

Table 3 shows eleven models addressed in publications over the last 10 years and their respective references. The model acronyms and the date of their first publication are shown in parentheses. Table 3 illustrates that three of the eleven models identified by Jokela et al. in 2006 have been addressed by publications in the last decade (italicized in Table 3). Although the following questions discuss the features of these models in detail, we should highlight that most of the publications on new UMMs (not italicized in Table 3) are categorized as solution papers that explain the theory underlying the model. An exception is the OS-UMM, which also reports a validation and evaluation of the model. The publications addressing models created before 2006 (italicized in Table 3) are mainly categorized as evaluation papers.

Table 3. UMMs addressed by publications over the last 10 years

UMM	Reference
AgileUX Model (Agile UX, 2014)	S3
UX Maturity Model(UX-MM, 2014)	S4
AUCDI Maturity Model (AUCDI-MM, 2013)	S2
Open Source Usability Maturity Model (OS-UMM, 2011)	S5
Health Usability Maturity Model (HU-MM, 2011)	S6
Corporate UX Maturity Model (CUX,MM, 2009)	S7
KESSU 2.2 (KESSU 2.2, 2007)	S8
Corporate Usability Maturity (CUM, 2006)	S9, S10, S11
ISO 18529 + ISO 15504 (ISO 18529 + 15504, 2000)	S13, S14, S16
Usability Maturity Model - Processes (UMM-P, 1998)	S15
Usability Maturity Model - HCS (UMM-HCS, 1998)	S12, S17

4.2 RQ2. General Features of UMMs

RQ2 aims to provide an overview of UMMs. Therefore, Table 4 summarizes a set of general characteristics of the eleven (11) models identified in our study according to the criteria used in [S1]. None of the models define the time required to achieve maturity, except for CUM. CUM's author, Nielsen, states that it takes 40 years to reach usability

Table 4. General features of UMMs

Model	Time Maturity	Target audience	Domain	Maturity Levels	Dimensions
Agile UX	Undefined	Mgmt and tech	Scrum	Not defined	Six dimensions, e.g., integration of user experience (UX) practices on development and budget for UX
UX-MM	Undefined	Mgmt	Any domain	Stage 1: Beginning to Stage 5: Exceptional	Six indicators, e.g., integration of UX and corporate processes and design thinking to drive UX
AUCDI-MM	Undefined	Mgmt and tech	Agile Org.	Level 0: Not Possible to Level 5: Cont Improvement	Four dimensions, e.g., agile-user centred design (A-UCD) integration process and UCD continuous improvement
OS-UMM	Undefined	Tech	OS Org.	Level 1: Preliminary to Level 5: Institutionalized	Four dimensions, e.g., usability assessment and usability methodology
HU-MM	Undefined	Mgmt	Health Org.	Phase 1: Unrecognized to Phase 5. Strategic	Five elements, e.g., process and infrastructure for usability
CUX-MM	Undefined	Mgmt	Any domain	Level 0: Initial to Level 4: Driven Corporation	Undefined
KESSU 2.2	Undefined	Tech	Any domain	Level 0: Incomplete to Level 5 Optimizing	Seven activities, e.g., identification of user task design and usability feedback
CUM	40 years	Mgmt	Any domain	Stage 1: Hostility to Stage 8: User-Driven Corp.	Five dimensions e.g., management attitude towards usability and strategic usability
ISO 18529 + 15504	Undefined	Mgmt and tech	Any domain	Level 0: Incomplete to Level 5: Optimizing	Seven processes, e.g., plan the HCD process and facilitate the human-system implementation
UMM-P	Undefined	Mgmt and tech	Any domain	Level 0: Incomplete to Level 5: Optimizing	Same dimensions as ISO 18529 + 15504
UMM-HCS	Undefined	Mgmt and tech	Any domain	Level X: Unrecognized to Level E: Institutionalized	Set of practices, e.g. user focus and human centred improvement of organization

maturity [S10]. Such a long time period may not be appealing for a software development organization keen for results. We should note, however, that it, in general, it is not possible to specify accurate times for improvement processes as many organization-dependent factors have a bearing on such processes. These factors include the readiness of the organization, the existence of effective processes and infrastructure to support a programme, and the skills and knowledge of the organization's people [11].

As regards the targeted audience, the results of the evaluation for four models will be mainly useful to management. For example, the UX-MM model focuses on indicators like UX expertise and resources or leadership and culture in the company. On the other hand, two models focus on technology. For instance, KESSU sets out to evaluate the performance of different usability activities conducted by the development team. Finally, the other models combine management and technology issues.

As regards the model application domain, most are generally applicable, that is, can be applied in any type of organization. However, two models are for very specific domains. HU-MM was developed in response to usability problems detected in health-related products. In addition, the OS-UMM model was developed for open source models (OSS). Finally, two of the models were specially designed for organizations enacting an agile development approach (AGILEUX and AUCDI-MM).

Another key feature is the number of levels or stages to achieve maturity in usability. According to Fraser et al. [12], a model usually defines up to six maturity levels. Most of the retrieved models are within this range, except KESSU and CUM with seven and eight stages, respectively. Still, this is not a major deviation. On the other hand, the information reported in the publication that we retrieved about the AGILEUX model is partial, as it only describes level 2 and does not refer to the total number of levels to be considered.

Finally, all the models, except CUX-MM, define areas, dimensions or criteria (depending on the model) that identify key structural elements in the field of usability. They are used to ascertain the usability maturity within an organization. The results of the evaluation of these areas illustrate maturity as a whole and separately for each of the evaluated areas or dimensions. Table 4 shows examples of these dimensions for the different UMMs.

4.3   RQ3. UMM Design Features

The design or structural features of a maturity model are used to describe the form and organization of the model. As already mentioned, we use the design attributes proposed by Mettler et al. [13] for maturity models applied to the information system field. Additionally, the values of some of these criteria were complemented by other research as mentioned below:

- Maturity concept defines the approach of the model:
  - Process maturity, that is, the extent to which a specified process is specifically defined, managed or controlled.
  - Object maturity, that is, the extent to which a particular object, for example, a software product reaches a predefined level of sophistication.

- Workforce maturity, that is, how proficient a team of people are at building knowledge and improving skills.
- Composition is, according to Fraser et al. [12], divided into three types:
  - Maturity grids usually have a narrative text describing the activities for each maturity level; their design complexity is moderate.
  - Likert-like questionnaires aim to rate specified statements on good practices at different maturity levels.
  - CMM-type models have a more formal architecture and are more complex because a broad spectrum of scales and subscales should be implemented to evaluate maturity.
- Reliability defines two categories:
  - Validated: a model can be validated, qualitatively by means of case studies or using quantitative questionnaires.
  - Verified: a quite accurate conceptual description and specification of the model is given without evidence of its practical use.

If there is no detailed information in this respect, the model is catalogued as "Not fully described".

- Mutability defines two categories:
  - Form refers to whether the model accounts for changes in the description of maturity levels and requirements in order to assure model standardization.
  - Operation refers to changes defined by the model on how maturity is measured at each stage.

Table 5 is a summary of design features of the analyzed UMMs. In this case, all the models are oriented to the usable software construction process.

Table 5. Maturity models design features

Model	Maturity	Composition	Reliability	Mutability
AgileUX	Process	CMM family	Verified	Not mentioned
UX-MM	Process	Maturity grid	Not fully described	Not mentioned
AUCDI-MM	Process	Likert-like	Verified	Not mentioned
OS-UMM	Process	Likert-like	Validated, case studies	Not mentioned
HU-MM	Process	Likert-like	Verified	Not mentioned
CUX-MM	Process	Not mentioned	Not fully described	Not mentioned
KESSU 2.2	Process	Likert-like	Verified	Not mentioned
CUM	Process	Maturity grid	Validated, case studies	Not mentioned
ISO 18529 + 15504	Process	Likert-like	Validated, case studies	Not mentioned
UMM-P	Process	Likert-like	Validated, case studies	Not mentioned
UMM-HCS	Process	Likert-like	Validated, case studies	Not mentioned

As regards composition, six models have a Likert-like composition, where the model authors select the scoring scheme at their discretion. Without a clear description, however, these scoring schemes can be confusing, ambiguous, and lead to mistaken

results. In this respect, Salah et al. [S17] claim that the description provided for the UMM-HCS scoring scheme is unsatisfactory. On the other hand, the UX-MM and CUM models have maturity grid composition. Therefore, the result of the evaluation largely depends on how the evaluator interprets the model. AGILEUX is based on the CMM model, but reports only information for maturity level 2. Finally, we were unable to gather enough information from the literature to determine the composition of CUX-MM. Note that the model's composition type is not necessarily a strength or weakness a priori; it depends on what facilities the evaluator is given for applying the respective model.

With respect to the reliability attribute, we found that there is evidence about the use of five out of the eleven models based on case studies. According to Mettler et al.'s terminology, therefore, five models have been validated. There is no empirical evidence for the other six models. Our study did not retrieve any papers containing evidence about the UMM-P model. However, Jokela et al. [S1] pointed out that several case studies were conducted prior to 2006, albeit with contradictory results. On this ground, our study considers this model to have been validated. Note that the fact that model has been validated does not necessarily mean that the results of the validation were successful. On the other hand, Table 5 classifies four models (AgileUX, AUCDI-MM, HU-MM and KESSU 2.2) as verified. These models have an accurate conceptual specification. The conceptual accuracy regarding the UX-MM and CUX-MM models was not found to be good enough in the retrieved literature. On this ground, they have been classified as not fully described.

4.4 RQ4. UMM Use Features

RQ4 is related to the practical application of the model. The attributes identified by Mettler et al. in this respect are complemented with others also provided in literature as follows:

- The method of application defines who applies the model. This can be classified as a self-assessment, or a third-party assisted assessment.
- The type of support to which the model user has access. Three options are given for this attribute: (1) the user is not given any support material; (2) the user is offered a textual description about how to conduct the evaluation; (3) the user is offered a software tool to conduct the evaluation.
- The Purpose of use, defined by De Bruin et al. [14] as:
  - Descriptive: the purpose is to evaluate the current status of the organization.
  - Prescriptive: the purpose, apart from evaluating the organization's current status, is to suggest improvement guidelines in order to progress to the next maturity level. According to Pöppelbuß et al. [17], maturity models claiming to serve a prescriptive purpose of use must provide at least: (1) a set of improvement measurements and recommendations; (2) a decision calculus to help to evaluate different alternatives; and (3) a procedure on how to specify and adapt the improvement measures. In our study, the models that comply with all three characteristics are catalogued as fully prescriptive, whereas models meeting at least one will be classed partially prescriptive. Additionally, according to Mettler

et al., the improvement recommendations may be explicit, that is, detail exactly what to do to improve an activity or process, or implicit that is, they are embedded in other general and non-specific comments.

Table 6 summarizes the characteristics of UMM usage. As regards the method of application, five models were identified as self-assessment models. The authors of OS-UMM and HUMM clearly state that these models are self-assessment ones. AUCDI-MM and KESSU are said to have been designed for a non-specialized audience (that is, evaluators) and should not consume too much time or external resources. Additionally, UMM-HCS [16] was used by Salah et al. as a self-assessment method [S17]. The authors of UMM-P state that their guides were designed for expert personnel like process improvement consultants. Although no mention is made of the application method for the ISO 18529 + ISO 15504 model, it is, according to [S16], a complex model, and the formal use of ISO 18529 is a job for a professional. On this ground, it has also been classified as a third-party assessment model. We were unable to establish the method of application of the other models from the retrieved information.

With regard to the type of model application support to which the user has access, HU-MM is the only model offering a software tool to conduct the evaluation. As shown in Table 6, another five models provide a narrative description of the activities to be evaluated, an explanation of the scoring scheme and a recording form. We did not find any references to possible evaluator support material for the other models.

Finally, as regards the purpose of use, we found that eight of the models are descriptive. On the other hand, three models are classed as partially prescriptive since provide a set of improvement measures. Note, however, that the recommendations are implicit. The HU-MM merely mentions that it will offer some suggestions. As it does not outline these recommendations in the published document, it was catalogued as descriptive.

Table 6. Maturity model use features

Model	Method of Application	Support of Application	Purpose of use
AgileUX	Not mentioned	None offered	Part. prescriptive, implicit
UX-MM	Not mentioned	None offered	Part. prescriptive, implicit
AUCDI-MM	Self-assessment	Textual description	Descriptive
OS-UMM	Self-assessment	None offered	Descriptive
HU-MM	Self-assessment	Software tool	Descriptive
CUX-MM	Not mentioned	None offered	Descriptive
KESSU 2.2	Self-Assessment	Textual description	Descriptive
CUM	Not mentioned	None offered	Part. prescriptive, implicit
ISO 18529 + 15504	Third party assisted	Textual description	Descriptive
UMM-P	Third party assisted	Textual description	Descriptive
UMM-HCS	Self-assessment	Textual description	Descriptive

5   Discussion and Conclusions

In this work, we aim to characterize the UMMs that have been researched in the literature over the last decade. Three of the eleven retrieved UMMs were designed prior to 2006, whereas eight new UMMs have been created since 2006.

Generally, one of the differences detected between the models identified more than one decade ago and the more recent ones is that, new UMMs have been proposed for specific contexts like agile developments or open software over the last decade. This is an interesting development, as it may result in a more efficient evaluation targeting the specific features of such domains.

Several points must be addressed on UMMs. At first, from a practical point of view is their reliability, that is, whether there is evidence about their application in real environments. In this respect, our study is consistent with earlier studies highlighting that the cross-checking of maturity models is insufficient. This study has found information on only five empirically tested models (36%). Model checking was qualitative based on case studies, which is consistent with Wendler's and Jokela et al.'s findings. In our study, ISO 18529 + ISO 15504 is the model for which there is most empirical evidence. However, as discussed in Sect. 4.4, the model is complex, and the assessment has to be made by experts in maturity models. This can, according to [S16], be a major drawback for the practical application of the model. This model is followed by UMM-HCS and CUM, although we identified some deficiencies in CUM scoring or inconsistencies in CUM terminology.

Another important discussion point is the support provided by the models for evaluators. In our study, we have found that five out of the 11 retrieved models (45%) do not offer specific guidance for identifying the usability maturity levels in an organization. The other six models have a narrative description of how to perform this evaluation. Upon evaluation, however some were regarded as hard to interpret. On the other hand, we identified only one software tool supporting evaluation for the HU-MM model. Although there is no guarantee of model application being objective, since this depends on the quality of the material, any support material or even a support tool for evaluation is better than none. Briefly, in this regard, our study pinpointed the same weakness already identified by Jokela et al. [S1] for pre-2006 models, where 46% of the identified models offered no specific guidance to give practitioners insight into how to apply the models.

On the other hand, most models studied serve the purpose of description, that is, output a view of the company's usability status. Our study did not retrieve any fully prescriptive models; however, we did identify three partially prescriptive models with implicit recommendations or improvement practices. Although the studies by Wendler et al. [8] and Jokela et al. [S1] do not refer to the purpose of use, it is useful for establishing which models not only offer information on the organization's usability level but take a step further into practice. Although by no means straightforward, prescriptive models enabling a company to move to the next usability maturity level is an important research issue to cover to promote an effective integration of usability practices into the development process.

Another striking finding is that none of the models refer to the mutability feature, thereby accounting for the possible adaption of the model to new usability practices or process changes. This is also an important feature as software development is a live process and new techniques and practices should emerge as part of continuous improvement. However, we think that mutability would be a desirable feature once the above weaknesses related mainly to model validation, support and improvement recommendations have been resolved. Note that, even if these constraints are overcome, it may not be easy to state that one particular model is better than others, basically because the choice of model is dependent on the features and priorities of each organization. For example, the ISO 18529 + ISO 15504 model has advanced design and use features, but it is a complex model and would not be suitable for application in a small organization.

In summary, as discussed throughout this paper, the field of UMMs cannot yet be considered mature, even though the first UMMs date back over 20 years. Our research aims to contribute to building cumulative research on UMMs as suggested by Jokela et al. in 2006. Although it is not easy to offer practitioners clear recommendations on the best UMMs, the characterization outlined here is a potential decision-making aid. From a research point of view, our characterization is based mainly on criteria already defined in literature for analysing maturity models. Therefore, its application to UMM provides a more robust analysis complementing previous research. On the other hand, we have highlighted open issues and opportunities for research to bring forward the area of UMMs. Mature UMMs will contribute to improve the integration of usability and user experience techniques in the software development process.

Appendix A: Mapping Studies

[S1] Jokela, T., et al.: A survey of usability capability maturity models: implications for practice and research. Behaviour & Information Technology. 25(3), 263–282 (2006)

[S2] Salah, D., Paige, R., & Cairns, P.: A Maturity Model for Integrating Agile Processes and User Centred Design. In: International Conference on Software Process Improvement and Capability Determination. Springer International Publishing (2016)

[S3] Peres, A.L., et al.: AGILEUX model: towards a reference model on integrating UX in developing software using agile methodologies. In: 2014 Agile Conference, AGILE 2014, pp. 61–63 (2014)

[S4] Chapman, L., Plewes, S.: A UX Maturity Model: Effective Introduction of UX into Organizations. In: International Conference of Design, User Experience, and Usability, pp. 12–22. Springer International Publishing (2014)

[S5] Raza, A., Capretz, L.-F., Ahmed, F.: An Open Source Usability Maturity Model (OS-UMM). Computers in Human Behavior. 28(4), pp. 1109–1121 (2012)

[S6]   HIMSS Usability Task Force.: Promoting usability in Health Organisations: Initial Steps and Progress towards a Healthcare Usability Maturity Model. Health Information and Management Systems Society (2011)

[S7]   Van, S.: Corporate User-Experience Maturity Model. In: Human Centered De-sign, pp. 635–639. Springer, Berlin Heidelberg (2009)

[S8]   Jokela, T.: Characterizations, requirements, and activities of user-centered de-sign—the KESSU 2.2 model. In: Maturing Usability, pp. 168–196. Springer, London (2008)

[S9]   Nielsen, J.: Nielsen's Alertbox: Corporate Usability Maturity: Stages 1–4 (2006)

[S10]   Nielsen, J.: Nielsen's Alertbox: Corporate Usability Maturity: Stages 5–8 (2006)

[S11]   Salah, D., Paige, R., Cairns, P.: Integrating agile development processes and user centred design- a place for usability maturity models? In: HCSE 2014. LNCS, vol. 8742, pp. 108–125. Springer, Heidelberg (2014) In: Sauer, S., Bogdan, C., Forbrig, P., Bernhaupt, R., Winckler, M. (eds.) HCSE 2014. LNCS, vol. 8742, pp. 108–125. (2014)

[S12]   Ashley, J., Desmond, K.: Usability maturity: a case study in planning and de-signing an enterprise application suite. In: Human Centered Design, pp. 579–584. Springer, Berlin Heidelberg (2009)

[S13]   Bevan, N.: Comparison of Case Studies of Usability Maturity Assessment and Process Improvement. In: CHI 2013 Conference, pp. 54–59 (2013)

[S14]   Rodríguez, A.: Extending OpenUP to Conform with the ISO Usability Maturity Model. In: Human-Centered Software Engineering, pp. 90–107. Springer, (2014)

[S15]   Joshi, A., Gupta, S.: Usability in India. Global usability, pp. 153–168. Springer (2011)

[S16]   Jokela, T., Kantola, N.: Name of the method: ISO 18529 (+ISO 15504). R3UEMs: Review, Report and Refine Usability Evaluation Methods, pp. 106–108 (2007)

[S17]   Salah, D., Paige, P., Cairns, P.: Observations on utilizing usability maturity model-human centrdness scale in integrating agile development processes and user-centred design. In: Rout, T., O'Connor, R.V., Dorling, A. (eds.) SPICE 2015. Communications in Computer and Information Science, vol. 526, pp. 159–173. Springer, Switzerland (2015)

References

1. ISO, International Software Quality Standard, ISO/IEC 25010. Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE) (2011)

2. Hoo, M.H., Jaafar, A.: Usability in practice: perception and practicality of management and practitioners. In: International Conference on Pattern Analysis and Intelligent Robotics (ICPAIR), vol. 2, pp. 211–216 (2011)

3. Nielsen, J., et al.: Return on Investment (ROI) for Usability, 4th edn. Nielsen Norman Group, Fremont (2008)

4. Bornoe, N., Stage, J.: Usability Engineering in the Wild: How Do Practitioners Integrate Usability Engineering in Software Development? In: Sauer, S., Bogdan, C., Forbrig, P., Bernhaupt, R., Winckler, M. (eds.) HCSE 2014. LNCS, vol. 8742, pp. 199–216. Springer, Heidelberg (2014). doi:10.1007/978-3-662-44811-3_12
5. Becker, J., Knackstedt, R., Pöppelbuß, D.: Developing maturity models for IT management. Bus. Inf. Syst. Eng. **1**(3), 213–222 (2009)
6. Wendler, R.: The maturity of maturity model research: a systematic mapping study. Inf. Softw. Technol. **54**(12), 1317–1339 (2012)
7. Petersen, K., et al.: Systematic mapping studies in software engineering. In: 12th International Conference on Evaluation and Assessment in Software Engineering, vol. 17 (1) (2008)
8. Hox, J.J., Boeije, H.R.: Data collection, primary vs. secondary. In: Encyclopedia of Social Measurement, vol. 1, pp. 593–599 (2005)
9. Wieringa, R., et al.: Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. Requirements Eng. **11**, 102–107 (2006)
10. Condori, N., et al.: A systematic mapping study on empirical evaluation of software requirements specifications techniques. In: 3rd International Symposium on Empirical Software Engineering and Measurement, pp. 502–505. IEEE Computer Society (2009)
11. Cmmifaq. http://www.cmmifaq.info
12. Fraser, P., Moultrie, J., Gregory, M.: The use of maturity models/grids as a tool in assessing product development capability. In: Engineering Management Conference, pp. 244–249. IEEE International (2002)
13. Mettler, T., Rohner, P., Winter, R.: Towards a classification of maturity models in information systems. In: Management of the interconnected world, pp. 333–340. Physica-Verlag HD (2010)
14. De Bruin, T., et al.: Understanding the main phases of developing a maturity assessment model. In: 16th Australasian conference on information systems. Sydney (2005)
15. Pöppelbuß, J., Röglinger, M.: What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management. In: 19th European Conference on Information Systems. Finland (2011)
16. Earthy, J.: Usability Maturity Model: Human Centredness Scale. INUSE Project deliverable D5.1.4(s). Version 1.2., Lloyd's Register of Shipping, London IE2016 INUSE Deliverable D5.1.4s (1998)
17. Petersen, K., Vakkalanka, S., Kuzniarz, L.: Guidelines for conducting systematic mapping studies in software engineering: an update. Inf. Soft. Technol. **64**, 1–18 (2015)