Challenges in Implementing CMMI® High Maturity: Lessons Learnt and Recommendations

Isabel Lopes Margarido*, João Pascoal Faria, Marco Vieira, Raul Moreira Vidal

* isabel.margarido@fe.up.pt

Abstract:

CMMI® allows organisations to improve the quality of their products and customer satisfaction; reduce cost, schedule and rework; and make their processes more predictable. However, this is not always the case, as there are differences in performance between CMMI organisations, depending not only on the context of the business, projects and team but also on the methodologies used in the implementation of the model practices. To better understand the difficulties of implementing CMMI we conducted research on the challenges that organisations face when implementing CMMI. We did a literature review, analysed the data of an SEI survey to organisations being appraised at high maturity levels and conducted three case studies from organisations previously appraised at CMMI maturity level 5. Empirical evidence shows that the identified problems are common to the different sources of information. To prevent such problems in the future, we map the challenges that the organisations faced with a set of proven recommendations in a checklist, to be used by organisations when implementing or improving CMMI practices.

Keywords: Capability Maturity Model Integration, High Maturity, SCAMPI, Implementation Plan, Checklist, Metrics, Processes

Introduction

Organisations that implement the Capability Maturity Model Integration (CMMI®) typically improve performance in terms of quality of products and processes, schedule and costs (Goldenson, Gibson, and Ferguson 2004). Consequently, processes become more predictable and customer satisfaction increases. Still, not all stories are successful.

Implementing CMMI Maturity Levels (MLs) 4 and 5[1], also known as high maturity levels (HMLs), is particularly challenging. Due to their dependency on metrics HMLs rely on the quality of implementation of the ML 2 Process Area (PA) Measurement and Analysis (MA) (Chrissis, Konrad, and Shrum 2011). When implementing MA organisations define metrics, how they are collected, stored and analysed; such metrics must be meaningful to the organisation and aligned with its business goals.

Considering ML 4, to implement Organisational Process Performance (OPP) it is essential to establish the organisation standard processes, build process performance models (PPM) and establish process performance baselines (PPB). In the case of Quantitative Project Management (QPM) one ensures that projects establish and monitor quantitative goals based on organisation, client and team goals, and select sub-processes to use and monitor their performance. In ML5, Organisational Performance Management (OPM) is implemented to guarantee that the organisation performs towards meeting its goals. If problems, deviations or needs for improvement are detected Causal Analysis and Resolution (CAR) is used to analyse root causes of problems and eliminate them. OPM also includes necessary practices to select and pilot improvements, verify effects in processes performance before deployment, and regulate improvements deployment.

The Standard CMMI Appraisal Method for Process Improvement (SCAMPI℠) may be used to benchmark the maturity of a company in the CMMI model (Hollenbach and Smith 2002), by identifying process weaknesses, investigating and determining the degree of satisfaction of CMMI PAs Goals, and assigning maturity ratings (CMU/SEI 2011). As it is not cost and effort effective to appraise an entire organisation and all its projects, SCAMPI includes sampling rules, to ensure that the appraised subset of

[1] Staged representation.

business units and projects is representative of the entire organisation. The selection of the necessary evidences from each appraised business unit follows coverage rules that determine the level of affirmations and artefacts that need to be provided (CMU/SEI 2011). However, these rules do not assure that all organisation projects use the practices of the appraised ML.

A survey (Radice 2000) to understand what CMM level 4 and 5 companies used in the implementation of the model showed that practices were not clearly institutionalised. The Software Engineering Institute (SEI) concluded that some companies did not understand the statistical nature of CMMI level 4 and companies that achieved CMMI HML rating did not have a consensus on necessary characteristics of level 4 (Hollenbach and Smith 2002). (Charette, Dwinnell, and McGarry 2004) reported issues such as lack of capability, poor performance and/or lack of adherence to processes which were found in the application of CMMI. The US Department of Defence (DoD) (Schaeffer 2004) recognised that not all organisation programmes are appraised, so practices are not implemented organisation wide and baselines erode once a certain ML is achieved. In response to this problem (Pyster 2006) proposed a set of solutions, one of which is already covered in SCAMPI version 1.3: provide guidance on how to select representative samples and aggregate results from subordinate organisations.

In this paper we analyse problems and difficulties occurring in the implementation of CMMI, reported in the literature and in three maturity level 5 organisations. Based on such research, we present recommendations that could help to avoid those problems. We sustain problems and recommendations with the analysis of the data of a SEI survey. Finally, we compile our results in a checklist to support organisations in the implementation of CMMI.

This paper is organised in sections: "Literature Review", introduces problems and difficulties faced when implementing CMMI, SCAMPI limitations and the results of SEI surveys involving HML organisations; we do "Further Analysis of the HML Survey Data"; "Case Studies", reports problems found in three organisations appraised at CMMI level 5; in "Problems Analysis" we analyse the problems that are common to the literature review, SEI surveys and the case studies; "Recommendations" towards preventing such problems; discuss "Validity and Limits to the Generalisation" of our results; in "Conclusions" we present our conclusions and propose future research.

Literature Review

When implementing CMMI HMLs, organisations may find difficulties in evolving existent processes and implementing new ones, as it is necessary to move to statistical thinking and quantitative management (Takara, Bettin, and Toledo 2007). Statistical Process Control (SPC), based on Shewhart Control Charts, is useful to implement metrics programmes, e.g., to define PPB. (Florac, Carleton, and Barnard 2000)discuss the challenges of applying SPC to complex processes with several sub-processes or the entire software process. SEI compiled PPMs used by CMMI ML 5 appraised organisations (Stoddard II and Goldenson 2009).

To build PPMs and PPBs it is necessary to collect data. Metrics programs have failed to produce appropriate metrics to analyse performance and capabilities of organisation processes (Barcellos 2009). In other cases, a great number of metrics were defined but remained unimplemented and unanalysed in decision making (Monteiro and Oliveira 2011). (Kitchenham et al. 2006) analysed a CMMI level 5 corporation database, finding that metrics were collected but could not be correlated; consequently they were meaningless to upper management. The authors detailed how to build databases and store data in a useful way, and how to link collected metrics, development environment and business context.

When the MA system is complex, the support information system becomes essential to avoid errors and overhead. (Johnson et al. 2005) recommended the use of telemetry to improve software development management and a set of metrics. The tool sensors connected to all software development tools collect process and product data non-intrusively. Other authors showed concern regarding metrics ambiguity: they are expressed in natural language (Goulão 2008) and can have different definitions in the literature, specifications of tools and when used in practice (Breuker et al. 2009). Literature needs to clearly define software metrics, and practitioners should be aware of this when implementing the MA system.

Furthermore, inadequate metrics may result from improperly deploying the indicators to control projects (Hamon and Pinette 2010), and frequently the CMMI program management is ineffective and people misinterpret CMMI (Leeson 2009) (Table 1).

Table 1: Metrics deployment, program management and CMMI interpretation issues.

Type of Problem	List of Problems
Inadequate Metrics Deployment (Hamon and Pinette 2010)	Meaningless, useless and non goal-driven indicators
	Complicated indicators without triggers for actions
	Inexperienced implementers
	Complex solutions, hard to maintain
	Same indicators for all situations (ignoring specific needs of different projects)
	Too many indicators at the beginning
	Outdated measurement plans
	Return of investment of the metric ignored
	Bad quality data
	Processes not followed
CMMI Program Management Issues (Leeson 2009)	Senior management not involved in establishing objectives, policies and the need for processes
	Sponsor not playing its role and delegating authority
	Software Engineering Performance Group not managed
	Organisations focused on achieving a maturity level more than improving quality of products or services
CMMI Understanding Issues (Leeson 2009)	Lack global view of the model
	Do not understand the relationship between measurement and project monitoring
	Do not understand several ML2 Generic Practices
	Do not understand difference between capability and maturity level
	Misinterpret ML2 and ML3
	Forget HML failing to understand the end-picture and not seeing the direction they are taking at lower levels before moving to HML

All the aforementioned problems result of a deficient implementation of CMMI but they become more nefarious when they are not detected in appraisal. Why they are not detected?

The SCAMPI has to be time and cost effective and some of its characteristics can be seen as limitations. The appraisal team analyses the practices implementation and how people actually work, but the quality of their results is out of the SCAMPI scope (Armstrong et al. 2002). With such orientation malpractices may be missed by the appraisal team. It is assumed that SCAMPI results depend on the quality of the appraisal team; in fact, they reflect their knowledge, experience and skills (CMU/SEI 2011). The SCAMPI relies on organisation's honesty: it provides evidences and supports the choice of projects to appraise (Armstrong et al. 2002). Either the lead appraiser is very rigorous in projects choice and critique about evidences, or the organisation bias may affect the appraisal outcome.

In the appraisal only a small number of affirmations sustain CMMI practices; SCAMPI V1.3 coverage rules limit the number of affirmations. Additionally, not all programs of the organisations are analysed (Schaeffer 2004), only a small percentage – meaning that practices may not be institutionalised.

We consider that a single affirmation from a business unit, safe by coverage rules, can demonstrate that a practice is not used.

A key problem is that measuring organisations performance is outside SCAMPI scope. The assessment verifies if the techniques applied allow achieving CMMI goals (Masters et al. 2007). There are only two PAs where performance improvements are explicitly analysed: CAR, where the effect of implemented actions on process performance should be evaluated; and OPM, where the selection and deployment of incremental and innovative improvements should be analysed (CMU/SEI 2008).

What can organisations do to ensure they properly achieve HML? This question is indirectly answered in the reports of two surveys conducted by the SEI: TR2008 (Goldenson, McCurley, and Stoddard II 2008) and TR2010 (McCurley and Goldenson 2010). The surveys, regarding the use and effects of measurement and analysis in HML organisations, were focused on the value added by PPM and the results were considered comparable. In 2008 the respondents were the sponsor assisted by a delegate, or its delegate, from organisations appraised at CMMI HMLs, in 2009 similar questions were asked to lead appraisers of organisations pursuing HMLs.

The SEI analysis is relevant to organisations pursuing any ML of CMMI, including problems and good practices that may have helped the organisations to achieve HML. In one question organisations had to indicate their routine while using PPM. The 2009 respondents indicated their most common problem was the long time it takes to accumulate historical data, which some organisations remediated by doing real time sampling of processes when they hadn't prior data available. In both surveys, the respondents gave different importance to obstacles found in the implementation of PPM (Figure 1) – organisations should consider them when implementing CMMI.

To measure the strength of the relationship between two variables and accuracy of predicting the rank of the response the authors used the Goodman and Kruskal's gamma. In 2008 the gamma value between the quality of the managers training and their capability to understand PPM results was not very strong. However, that relation was stronger when the training was more formal. 80% of respondents considered that the builders and maintainers of PPM understood CMMI's definition of PPM and PPB very well or even better, but their perception of the circumstances under which PPM and PPB are useful was lower. The results improved in 2009, over 50% of appraisers considered that the builders and maintainers of PPM understood all concepts very well or extremely well.

Organisations reported difficulties in collecting data manually. Regarding the automated support for MA activities responses to the 2008 survey showed that the organisations used spreadsheets, automated data collection and management software and, less frequently, statistical packages, workflow automation or report preparation software. Automation was considered to have a moderately strong relation with the overall value of PPM.

There was also a strong relationship between the overall value of PPM and the following variables:

- Models with emphasis on healthy[2] ingredients, and models for purposes consistent with those ingredients;
- Diversity of models used to predict product quality and process performance;
- Use of statistical methods[3];
- Data quality and integrity checks;
- Use of simulation or optimisation methods[4].

[2] The authors compiled a list of "healthy ingredients" to be considered in process performance modelling – page 46, TR2010.

[3] Regression analysis for prediction, analysis of variance, SPC charts, designs of experiments.

[4] Monte Carlo simulation, discrete event simulation, Markov or Petri-net models, probabilistic models, neural networks, optimisation.

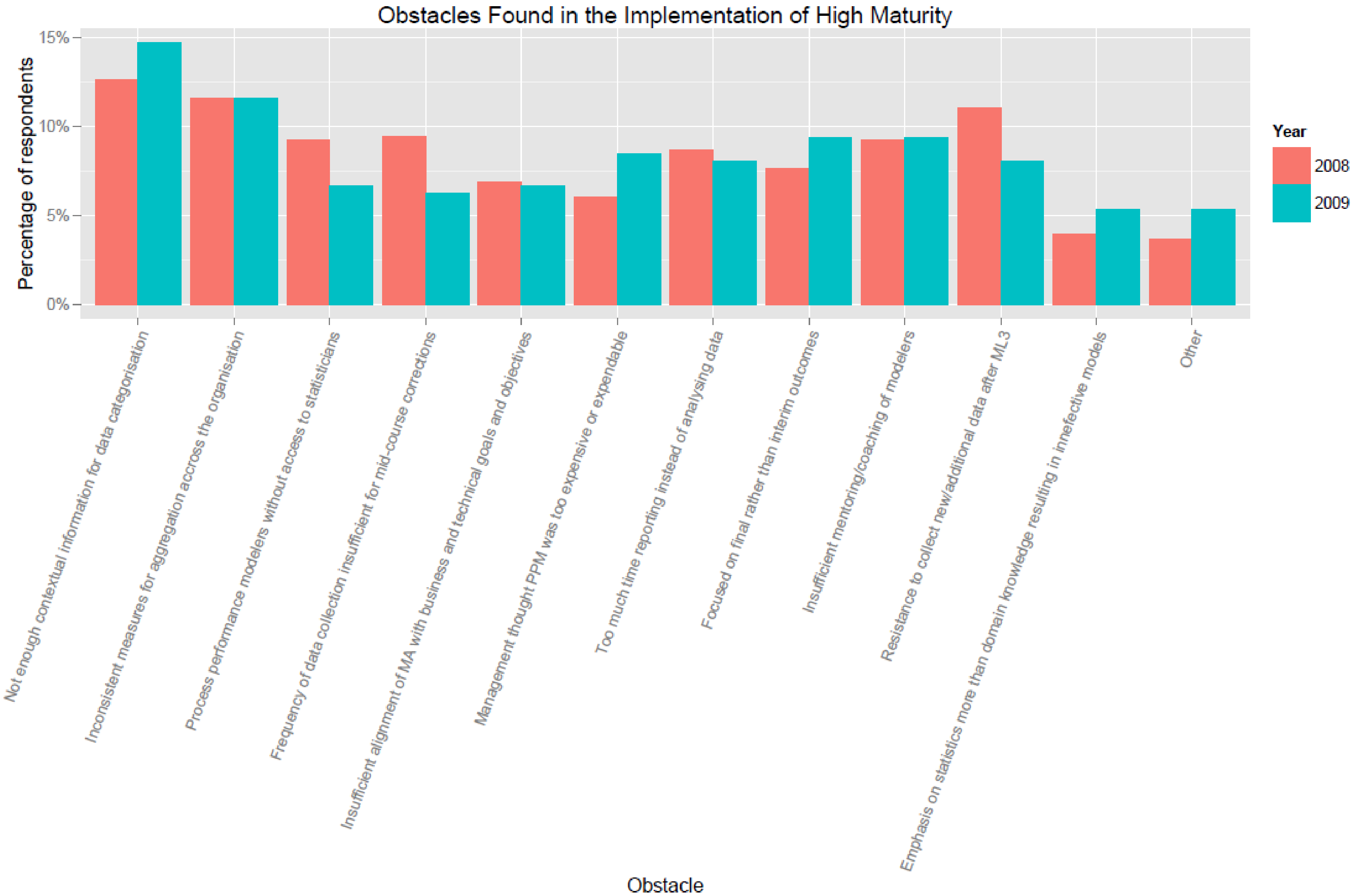

Figure 1: Obstacles identified by the organisations respondents found in the implementation of HML (TR2010).

When analysing the relations between the achievement of HML and certain practices some revealed differences between achieving and not achieving HML:

- All organisations with poor or fair documentation relative to process performance and quality measurement results failed to achieve high maturity, whilst most of the organisations with excellent and good documentation achieved HML;
- Using simulation/optimisation techniques had a strong relation with HML achievement. The relation with the number of such methods used and achieving HML was very high. Particularly, all organisations using two of those methods achieved HML;
- There was a very strong relation between achieving HML and, respectively: having models with emphasis on healthy ingredients, and the models for purposes consistent with those ingredients;
- All organisations that used statistical techniques substantially, achieved HML;
- The frequency of using PPM predictions in status and milestones reviews had a quite strong relation with the achievement of the target HML.

In general the gamma between variables was higher in the 2009 survey. (TR2010), justify the improvements from the 2008 to the 2009 surveys results: “There may be a trend over time” and/or “The perspectives of the sponsors or the appraisers are more accurate”. The results of both surveys indicate several improvements that HML organisations may consider to get full advantage of having PPM in place, but also show the obstacles that organisations that intend to implement HML practices may find.

Further Analysis of the Survey Data

We analysed the data of the 2009 survey to sustain problems and recommendations with empirical evidence. Regarding the recognition of the efforts of people involved in MA initiatives, even if the difference is not very significant, the graph (Figure 2) shows that a larger number of organisations that achieved HML did give promotions or monetary incentives. The difference seems to be more relevant when they were given to Project Engineers, Technical Staff and Project Managers, i.e. people working closer to the projects.

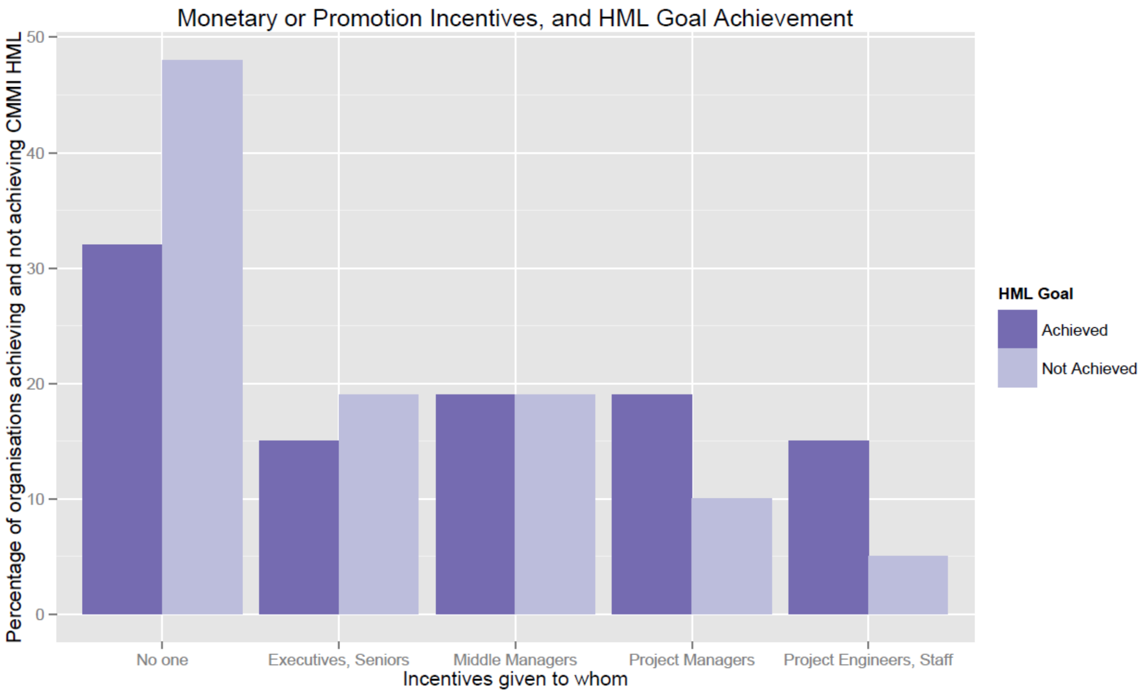

Figure 2: Relation between giving incentives to people who use use and improve MA, and the achievement of the HML goal.

We used the statistics described in Table 2 to understand the relation between achieving HML (V1) and doing a given practice. We tested the dependency of several variables and compared the groups of organisations that achieved HML and the ones that did not. In Table 3 we report the tests results of dependent variables (p-value < 0.05 in the Chi-square test) with different central tendency between groups (p-value < 0.05 in the Mann-Whitney test).

Table 2: Statistics and hypotheses that were tested.

Test	Description	Hypotheses
Levene	**Purpose:** Verify conditions to use the Mann-Whitney U test. The samples must have the same variance.	*H_0 (Null Hypothesis) – the two samples (Achieved and Not Achieved) have the same variance*
	If p-value < 0.05 we can reject the null hypothesis: the variance of the two samples are different. If p-value >= 0.05 keep the null hypothesis: the variances of the two samples are the same.	*H_1 (Alternative Hypothesis) – the two samples have different variance*
Mann-Whitney	**Purpose:** Verify if the two samples have similar median.	*H_0 – the two samples have similar median*
	If p-value < 0.05 we can reject the null hypothesis: the median of the two groups is different. If p-value >= 0.05 there is no difference between the groups.	*H_1 – median of Achieved > median of Not Achieved*
Chi-square	**Purpose:** Verify dependency between two variables.	*H_0 – the variables are independent*
	If p-value < 0.05 we can reject the null hypothesis: the variables are dependent. If p-value >= 0.05 the variables are independent.	*H_1 – the variables are dependent*

Table 3: HML 2009 survey – further data analysis. Results of the tests done with the groups of organisations that achieved and did not achieve HML and that were shown to have the same variance through the Levene test.

Variables	Levene	Mann-Whitney U	Chi-Square
V1: CMMI HML goal Achievement V2: How well managers understand PPM and PPB results	F = 0.0399 p-value=0.8423	W = 200.5 p-value = 1.44e-05	X-square = 20.64731 p-value = 0.0003719468
V1 V3: PPM and PPB creators understand the definition of PPM given by CMMI	F = 2.1284 p-value = 0.1489	W = 875 p-value = 7.75e-05	X-square = 3.271229 p-value = 0.0003911981
V1 V4: PPM and PPB creators understanding of the definition of PPM given by CMMI	F = 1.4462 p-value = 0.2333	W = 875 p-value = 7.75e-05	X-square = 20.53658 p-value = 0.0003911981
V1 V5: PPM and PPB creators understand the definition of PPB given by CMMI	F = 0.0484 p-value = 0.8264	W = 902.5 p-value = 1.711e-05	X-square = 19.64407 p-value = 0.0005870212
V1 V6: PPM and PPB creators understand when PPM are useful	F = 0.0082 p-value = 0.9281	W = 920 p-value = 7.735e-06	X-Square = 23.23521 p-value = 0.000113636
V1 V7: PPM and PPB creators understand when PPB are useful	F = 0.1445 p-value = 0.7051	W = 931.5 p-value = 4.198e-06	X-square = 24.84557 p-value = 5.403748e-05
V2 V4,5,6,7	F = 3.1665 p-value = 0.07963	W = 576.5 p-value = 0.02413	X-Square = 9.809072 p = 0.0202608
V2 V8: Availability of experts to work in PPM	F = 0.0095 p-value = 0.9227	W = 163 p-value = 2.094e-05	X- Square = 23.21129 p-value = 0.000114894
V1 V9: Distinguishing missing data from zeros	F = 0.8992 p-value = 0.3463	W = 855.5 p-value = 2.249e-05	X-square = 16.1262 p-value = 5.926e-05
V1 V10: Checking data precision and accuracy	F = 2.5641 p-value = 0.1139	W = 761 p-value = 0.00389	X-square = 7.0344 p-value = 0.007996

We verified that there is a relation between the understanding that the creators of PPM have of the CMMI intent (V4 to V7) and achieving HML. Figure 3 shows that a bigger percentage of organisations that achieved HML understood very well or extremely well the intent of CMMI. In both surveys the relation between understanding results of PPM and PPB, by Managers who use them (V2), and the achievement of HML was quite strong (Figure 4). We observed a relation between managers that understand the results better (V2) and creators that understand better the CMMI intent (V4 to V7) (Figure 5). A similar behaviour was verified with the availability of PPM experts to work (V8) (Figure 6).

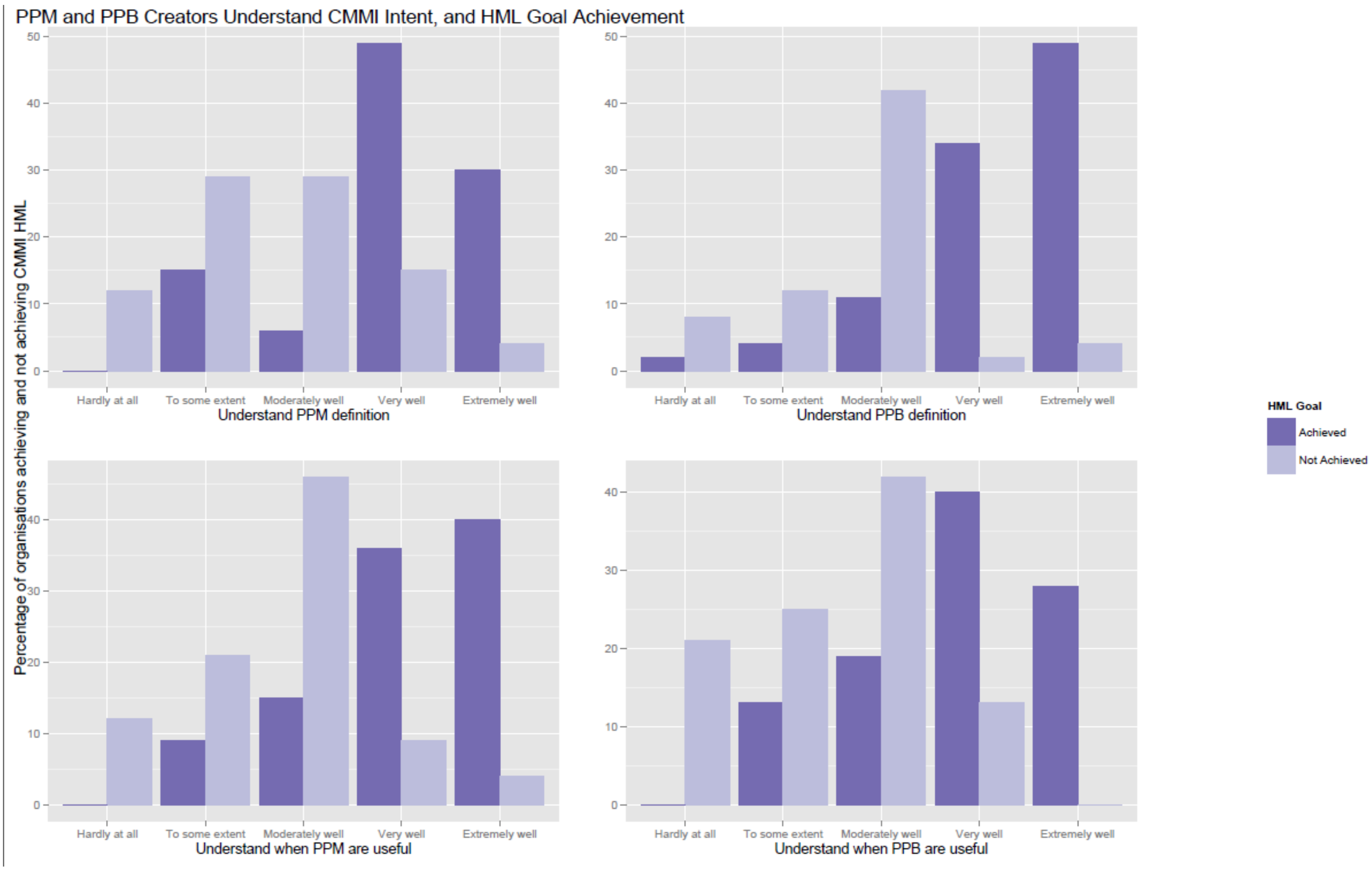

Figure 3: Relation between understanding the CMMI intent with PPM and PPB by their creators, and the achievement of the HML goal.

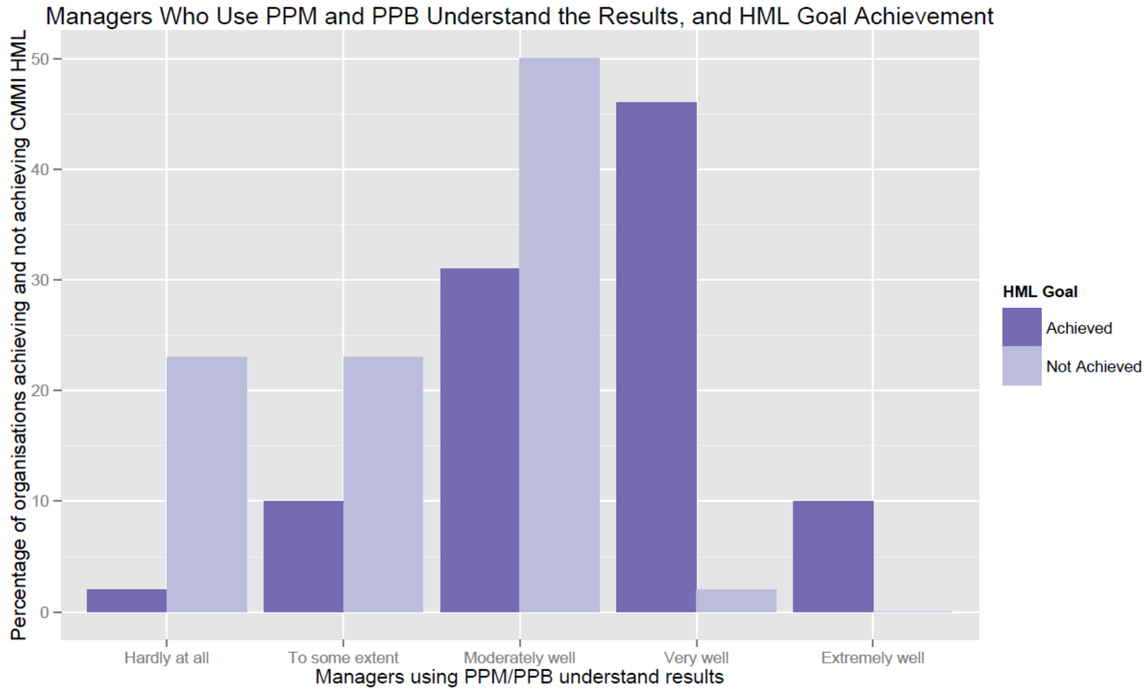

Figure 4: Relation between managers who use PPM and PPB understanding the obtained results, and the achievement of the HML goal.

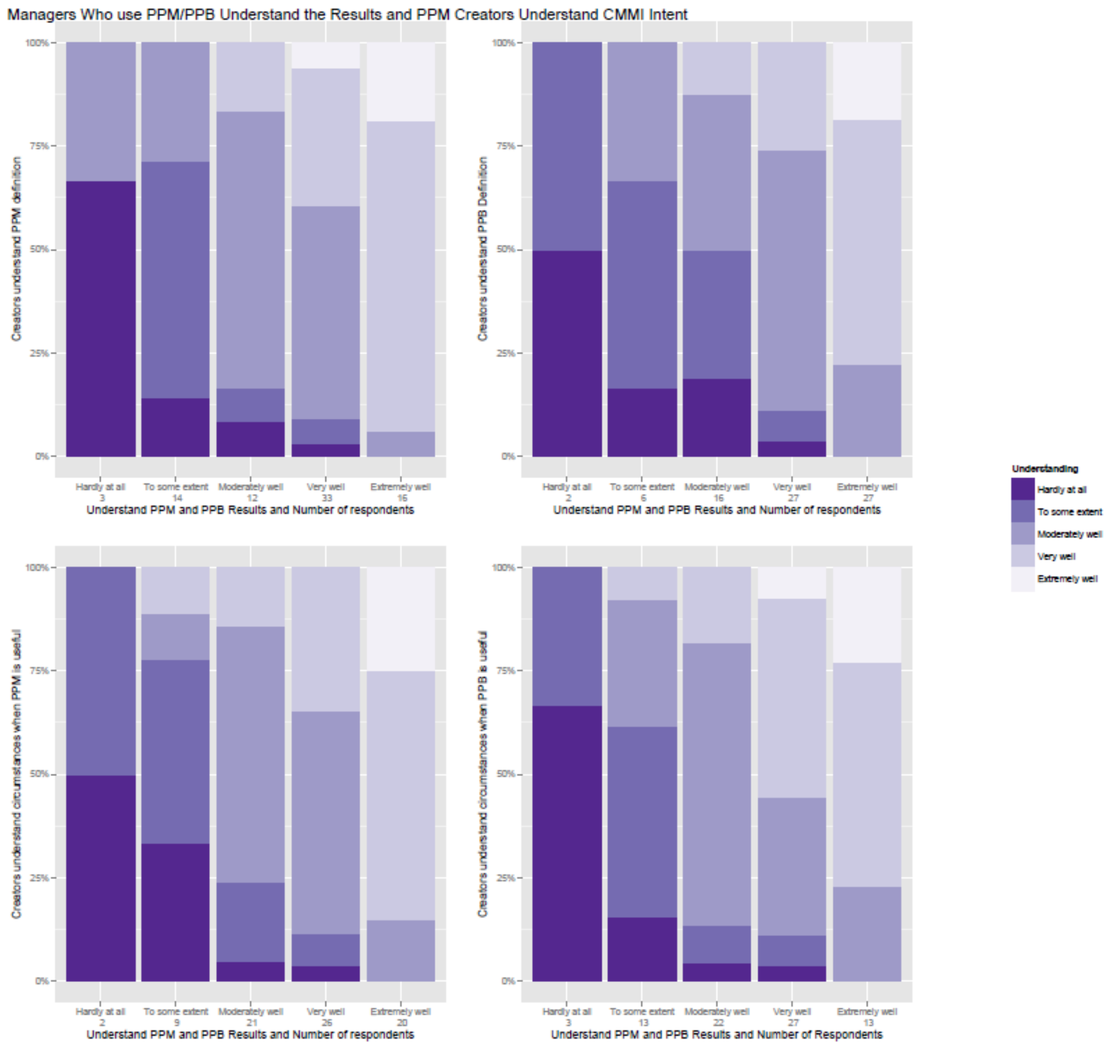

Figure 5: Relation between PPM and PPB creators understanding the CMMI intent and managers who use them understanding their results.

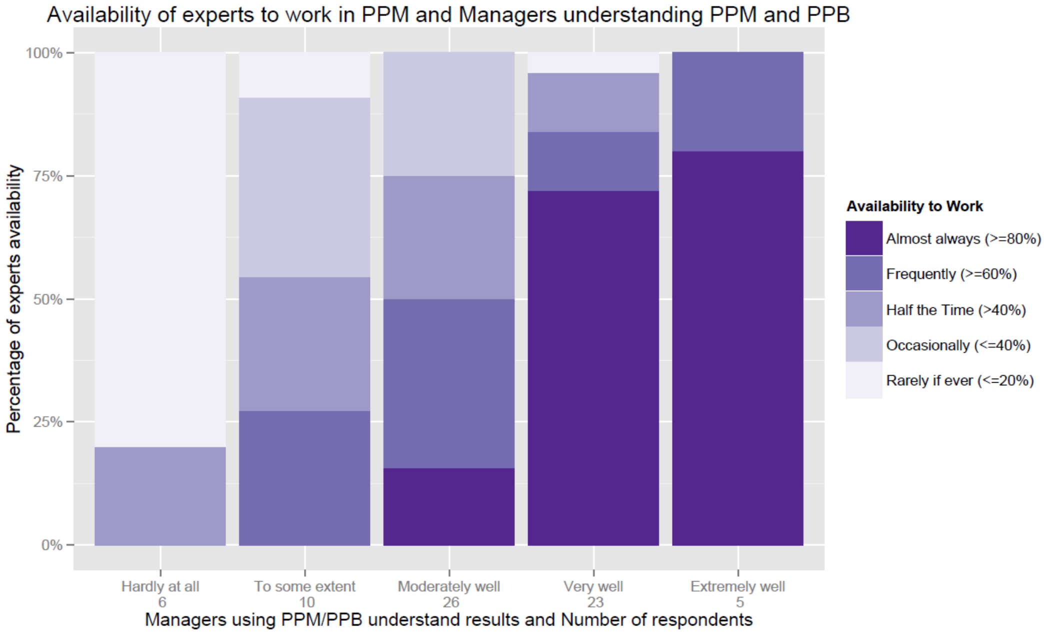

Figure 6: Relation between the availability of experts to work in PPM and managers who use them understanding their results.

The graphs in Figure 7 indicate that integrity data checks also seem to be related with the achievement of HML. The statistical tests support that distinguishing missing data from zeros and checking data precision and accuracy are related with achieving HML.

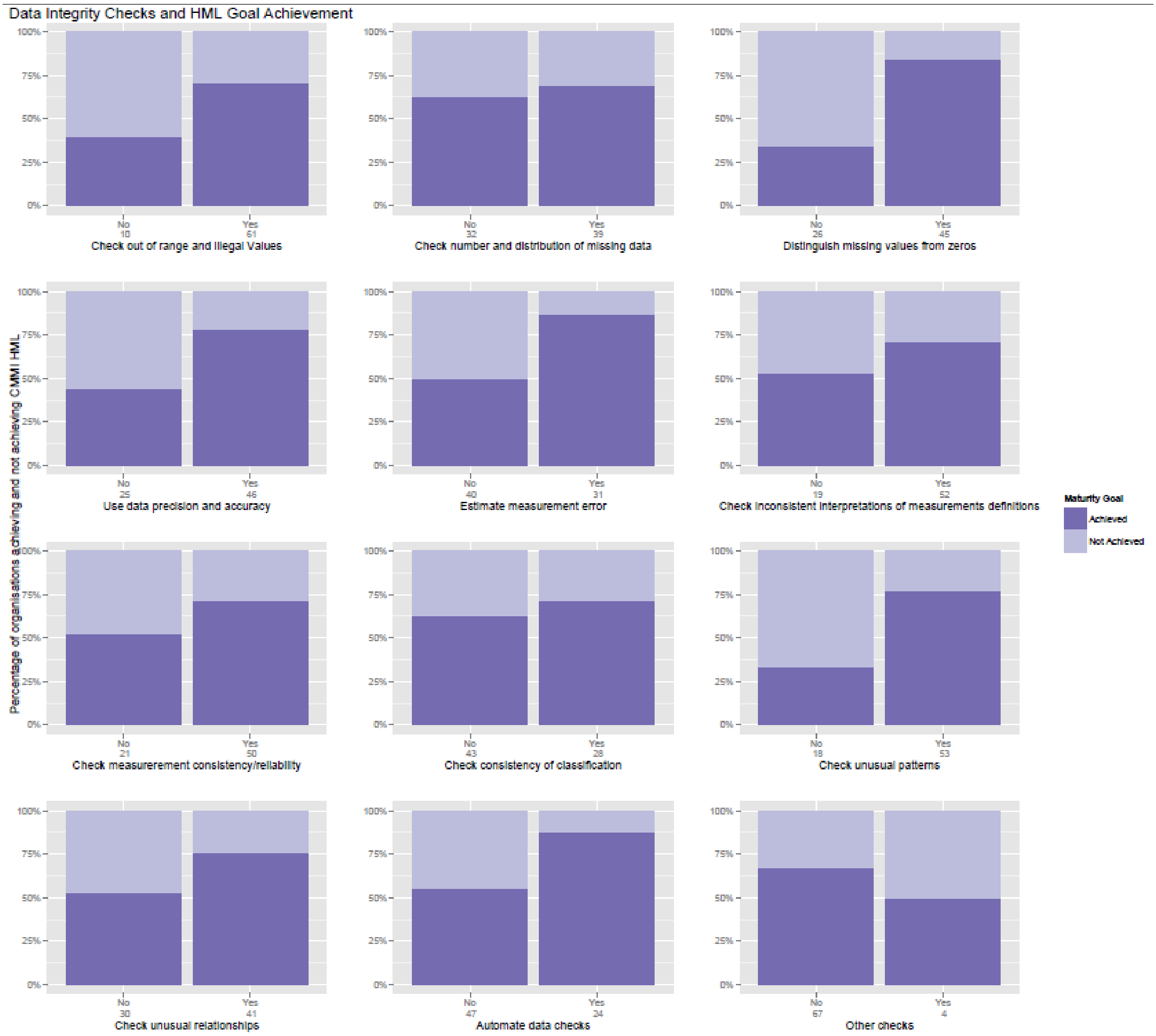

Figure 7: Relation between performing data integrity checks and the achievement of the CMMI HML goal.

With the analysis that we did we found that the following variables are related with the achievement of the desired HML:

- How well managers understand PPM and PPB;
- How well PPM and PPB creators understand the CMMI intent;
- Distinguish missing data from zeros;
- Check data precision and accuracy.

These relations reinforce the importance of doing proper integrity data checks to have meaningful and reliable PPM and PPB supporting high maturity PAs. Furthermore, to ensure that managers understand PPM and PPB results correctly, and consequently take appropriate actions, PPM and PPB creators must understand their CMMI meaning and usefulness, and experts must be available.

Case Studies

To verify the problems that a CMMI implementation may present, we conducted three case studies in multinational organisations that develop software (CI, CII and CIII) assessed at CMMI for Development ML 5, staged representation. Case studies on CI and CIII were conducted immediately after the appraisal. Our purpose was to identify real problems and difficulties in the implementation of CMMI and find recommendations to avoid them. We mainly focused on MA and HML, but also analysed other CMMI PAs. The research questions we intended to answer, which we considered when designing the case studies and analysing all data, were the following:

- What was the strategy to evolve to the new ML?
- What difficulties and problems occurred in the implementation of the new practices?
- What is the process definition?
- How was the process defined?
- How are people using the process?
- How are people collecting, analysing and interpreting process data?
- What is the impact of the new process on people's work?

Results

While analysing these case studies we detected some of the problems mentioned in the previous sections, which we identify with DS (Data of the Surveys) and with the author(s) initial(s).

Case Study I

In CI we interviewed the CMMI programme sponsor, posing direct questions and an open-end question to which the interviewee answered by narrating the story of the program. We had a similar interview with the program responsible. In both interviews we identified interviewees, projects and other documentation to analyse. We analysed the company Quality Management System (QMS), Information System and SCAMPI A repository. We also interviewed practices and tools implementers, and project teams, including the appraised ones (whose documentation we analysed).

Analysing CI data we found that the main problem stemmed from rapidly evolving to ML5 without giving enough time to have stable tools, processes, PPB and people behaviour.

Underestimate time to implement HML (DS) – the organisation re-planned the CMMI implementation programme several times until Six Sigma was introduced, allowing them to better understand HML demands.

Understand the statistical nature of level 4 (HS, TBT, DS) – Move to statistical thinking and quantitative management was their main challenge.

Dissemination problems (DS) – People noticed communication improvements. Nonetheless, the dissemination of information regarding processes and tools usage was not totally effective: some people still had difficulties to apply the new practices.

Lack of institutionalisation (R, S, CDM, HP) – Not all project teams were applying the new practices. This problem was also related to people behaviour and resistance to change.

Meaningless uncorrelated metrics (K, DS) – We found a case of metrics being misinterpreted due to lack of understanding of the context of one business area.

Metrics definition (collect and analyse data) (G, HP, B, DS) – People still had difficulties in collecting data in certain contexts and in their interpretation.

Metrics categorisation (DS) – The collection of data for high maturity had been occurring for a short period of time so the baselines were not stable enough. It was not possible to distinguish between different categories of data (for different markets, team experience, team sizes and project sizes), therefore the data were compiled in PPB categorised by technology only.

Baselines not applicable to all projects – PPBs were still unstable, or inadequate for all types of projects. Time to collect data was insufficient to gather information of different contexts and verify: *1)* if new metrics were needed; *2)* if there were differences in performance and in which contexts; *3)* if in certain circumstances the procedure to collect the data should be different.

Abusive elimination of outliers[5] – We found one situation when it was not perceived that the outlier occurred at least once in some projects. Some outliers are just indicators that the process improved its performance (Spirula member, 2010 personal communication).

Not all projects are measurable (DS) – Tools were not yet prepared to collect data in certain projects (maintenance, with several phases or outsourced), because their data structure was different from the standard projects (development). Besides, measurements specific to maintenance and outsource projects were not defined.

Tools setup and *Tools requirements* – The existing Information System evolved to support the new practices, but people were still detecting problems and requesting improvements, in tools and processes, resultant of using the tools in practice and in different projects contexts.

Overhead – To record time spent on tasks people manually filled a form per task, new metrics data collection was only partially automated. This was also due to the insufficient *Tools setup* time.

People behaviour – Changing mentality was a challenge; some people did not see value in new practices or stated that they were not applicable to their projects. It is difficult to convince people to report effort accurately: they normally report contract hours, not real effort, or do not report effort as they are finishing tasks, leaving the reporting until later.

Case Study II

CII is a business unit, located in several countries, that is part of CIIG, a CMMI level 5 organisation. In CII we interviewed the responsible for the CMMI programme, beginning with directed questions and finishing with descriptive questions regarding the story of the program. Afterwards we analysed CIIG QMS

[5] Data points that occur outside statistical limits of control of a process, indicating it is out of control; can result of an error on data collection or the circumstances in which they occurred can be completely characterised and are not expected to reoccur.

and Information System. Finally, we had a meeting with the CMMI program responsible and discussed our results and conclusions.

In CII we found several problems related to metrics; most limitations came from the fact that size was not being measured, and time spent on tasks stopped being accurately collected. Many of the identified problems were originated by the resistance to change and difficulty in presenting, to CIIG, metrics adequate for CII.

Copied processes – We found some processes in the CIIG QMS that were copies of CMMI, not reflecting the organisation's culture. That may have been one of the reasons for the problem *Multicultural Environment*.

Multicultural environment and *Impose processes* – CIIG acquired other companies and imposed its processes on them; consequently, their good practices, certain metrics and good visibility of processes were often lost. Another problem was that people from different cultures have different ways of working: in certain cultures orders are taken without question, while in others people need to understand the benefits of working in a certain way, otherwise they will resist change.

Lack of institutionalisation – In CIIG not all projects and business units performed at the same maturity level.

Metrics definition (collect and analyse data) – CIIG imposed KLOC (thousand lines of code) as the applicable size metric, which CII did not consider adequate to their types of projects.

Not all projects are measurable – Due to the problem above, many derived measures were not used by CII.

Baselines not applicable to all projects – CIIG had centralised PPBs not applicable to all business units' realities and projects, e.g., the development lifecycle phases had different durations depending on the business. Another problem was that productivity was measured considering a business day as unit of time but in some locations it had different duration in number of hours.

Effort estimates – While in CIIG effort estimation was based on their historical data of effort and size, CII estimates were based on expert judgment without using any tools or models.

People behaviour – CII workers stopped reporting effort accurately, only reporting contractual hours of work.

Case Study III

In CIII we interviewed a consultant involved in the appraisal of the organisation, i.e. a person who performed an actual observation on the case.

The main difficulty faced by CIII was to move to statistical thinking.

Underestimate time to implement HML – Implementation turned out to be more complex than anticipated and the programme took longer than planned.

Understand the statistical nature of level 4 – Changing mentality to HML was a significant shift because preparing and using the quantitative component takes time to mature. This problem may have been one of the causes of *underestimating time to implement HML*.

Introduction of HML forgetting ML 2 and 3 (L) – Occurs often and CIII was no exception.

First data collected were uncorrelated (HP) – Occurred at start, which implied conducting new data collection cycles and new searches for correlations. This may have been one of the causes of *underestimating time to implement HML* in CIII.

There are several common problems in the implementation of CMMI. We found them in the literature, organisations surveyed by SEI, and organisations of our case study. We analyse the intersections in the next section.

Problems Analysis

In Table 4 we signal where the problems we found occurred with “Yes”. 59% of the problems were shared either with other organisations or with the literature. The number of problems found in each organisation increased with the depth and insight provided by a more complete design of the case study. Nevertheless, we found two groups of problems common to two different groups of two organisations.

Table 4 Problems found in the case study organisations, the organisations surveyed by the SEI (DS) and the literature review (LR).

	Found In				
Problem	***CI***	***CII***	***CIII***	***LR***	***DS***
P1. Underestimate time to implement HML	Yes	Yes	Yes		Yes
P2. Introduce HML forgetting ML 2 and 3			Yes	Yes	
P3. Understand the statistical/quantitative nature of level 4	Yes	Yes	Yes	Yes	Yes
P4. Copied processes		Yes			
P5. Multicultural environments		Yes			
P6. Imposed processes		Yes			
P7. Dissemination Problem	Yes				Yes
P8. Lack of Institutionalisation	Yes	Yes		Yes	
P9. Meaningless Uncorrelated Metrics	Yes			Yes	Yes
P10. Metrics definition (collect and analyse data)	Yes	Yes		Yes	Yes
P11. First data collected were uncorrelated			Yes	Yes	
P12. Metrics Categorisation	Yes	Yes			Yes
P13. Baselines not applicable to all projects	Yes	Yes			
P14. Abusive elimination of outliers	Yes				
P15. Not all projects are measurable	Yes	Yes			Yes
P16. Effort Estimates		Yes			
P17. People Behaviour	Yes	Yes			Yes
P18. Tools Setup	Yes				
P19. Overhead	Yes				
P20. Tools Requirements	Yes				
P21.Complicated Indicators without triggers for actions				Yes	
P22. Inexperienced Implementers				Yes	
P23. Complex solutions hard to maintain				Yes	
P24. Out of date measurement plans				Yes	
P25. Return of investment of metrics ignored				Yes	
P26. Senior management not involved in establishing objectives, policies and the need for processes				Yes	Yes
P27. Sponsor not playing its role and delegating authority				Yes	
P28. Software Engineering Performance Group not managed				Yes	
P29. Organisations focused on achieving ML more than improving the quality of their products or services				Yes	
P30. CMMI not understood				Yes	Yes
P31. PPM focused on final rather interin outcomes					Yes
P32. PPM considered expensive and expendable by management					Yes
P33. Too much time reporting instead of analysing					Yes
P34. Frequency of data collection insufficient for mid-course corrections					Yes
P35.Trouble convincing management about value					Yes
Total:	14	12	4	16	15
Exclusive:	4	4	0	7	10
Shared:	10	8	4	9	5

Several problems found in CI were also detected in CII, four of them are related with metrics definition and usage and the other two are related with institutionalisation and people behaviour, respectively. Another two problems found in CI also occurred in CIII, all of them related with assuring entry conditions. CII was just a business unit of CIIG, who was rated ML5 for a long time, so we cannot verify if they faced similar entry conditions problems. However, we realised that the metrics problems found could be due to CII lack of understanding of the requirements for HML and statistical nature of ML4. We cannot even conclude that CIII did not face the metrics problems because we did not analyse their PPM, PPB, metrics definitions and usage in person.

37.5% of problems that we found in the literature were also detected in CI, CI and CIII, CI and CII, and CIII, respectively. 53.3% of problems found in the DS organisations were common to the ones found in our case study organisations, 16.7% of which were also found in the literature review.

Even if there are limits to the generalisation of our results the percentage of problems shared in more than one organisation/source indicates that they can occur when implementing HML, so organisations should be aware of them. In the next section we share a set of solutions that could help avoid the mentioned problems.

Recommendations

The demands of levels 2 and 3 should prepare organisations to adequately use measurement at higher levels, by monitoring appropriate metrics. Nonetheless, some of the problems identified in the previous section reflect a poor implementation of the MA PA, affecting the organisations results. Such problems become evident when implementing ML 4 because the correlation between variables and problems in the collected data, affect PPM and PPB. Besides, SCAMPI cannot appraise the entire organisation and does not analyse performance measures – if it did, it would become even more expensive. Hence, CMMI rating *per se* is not a guarantee of achieving expected performance results and organisations need to be aware that there are different methods that can be used on its implementation. Nevertheless, if some recommendations such as the ones we propose in this section are followed, CMMI implementation can be easier, and the problems discussed before can be avoided. Most of these recommendations are solutions used by the studied organisations to overcome their problems.

Table 5 summarises problems (P1 to P20) and recommendations (R1 to R36) that shall be considered when implementing CMMI, in particular ML 4 and 5, and can be used as a checklist. We also indicate the PA, Specific Practice (SP), Generic Goal (GG) or Generic Practice (GP), which are possibly affected.

Table 5 Problems found in the case studies and recommendations.

Category	Problem	Recommendations	Refs	PA/GG
Entry Conditions	*P1. Underestimate time to implement HML*, it takes long to accumulate meaningful historical data.	R1: Plan considering activities such as maturing levels, analysing and understanding HML, maturing PPB and PPM, collecting data repeatedly until meaningful performance indicators can be systematically obtained.	CI,CIII, DS	
	P2. Introduce HML forgetting ML 2 and 3	R2: Before moving to HML guarantee that ML 2 and 3 are mature and institutionalised.	CIII,L	
	P3. Understand the statistical/quantitative nature of level 4: Underestimation of time to change mentality from ML 3 to quantitative thinking, and time to implement ML 5. Insufficient involvement of management/stakeholders in models decisions. Ineffective models.	R3: Involve a statistician with experience in software and preferably on CMMI. R4: Introduce Six Sigma initiative. R5: Review goals and quantitative goals top down and bottom up when implementing CMMI.	CI,CIII, DS,HS, TBT L	MA OPP QPM
Process Definition and Implementation	*P4. Copied processes:* from CMMI.	R6: Processes shall reflect the culture of the organisation, not be a copy of the model imposed to the personnel. R7: Involve experts and process users in the definition of processes.	CII L	
	P5. Multicultural environments: people dealing differently with change.	R8: Interaction between business units to share processes and lessons learnt to design processes together.	CII	
	P6. Impose processes on acquired organisations with the loss of good practices.	R8, R9: Have goals specific to different business units, departments and projects, related to the organisation business goals. R10: Have indicators to monitor them at different report levels.	CII,16	
	P7. Dissemination problems: Difficulties in applying new practices, in particular in understanding how to collect, analyse and interpret metrics. Managers using PPM/PPB do not understand results and PPM modellers lacking mentoring/coaching and access to a statistician.	R11: Have commitment from the entire organisation: involve top management, middle management and the people who are actually doing the work. Have a sponsor. R12: Training shall include what to do, how to do, hands on, benefits and how can benefits be seen. Have different levels of training. Specialised training for sponsors and top management, process group and all roles that are affected by changes. Train top management on: sponsorship; goal setting; monitoring and rewarding (at different goals levels); on the process (understand it). R13: Coaching of projects and people (guiding and accompanying) and monitoring (from top management).	CI,DS L,H	GP2.5 GP2.6
	P8. Lack of institutionalisation: Not all projects used the new practices. Lack of adherence to processes or unused processes.	R14: Top management: set goals (when, who, what); include goals for gradual institutionalisation, monitor and reward. R13, R15: Metrics and processes definitions mature when used in practice, need time to define final versions.	CI,CII, R,S,CD M, HP	GG 2 GP2.5
Metrics Definition	*P9. Meaningless Uncorrelated Metrics:* Misinterpretation of metrics due to lack of context information; useless indicators and at times not aligned with business and technological goals.	R16: Use goal driven measurement, or equivalent, to establish quantitative goals. R17: Measures defined with a set of repeatable rules for collecting and unambiguously understand the data and what it represents. R18: Use different size measures according with the work product. R19: Identify different context that need to be associated with the metrics in order to adequately interpret them. R20: Do variables normalisation to ensure that metrics are usable in the entire organisation.	CI,DS,K PGF, FCB	MA SP1.4 MA SP2.2
	P10. Metrics definition (collect and analyse data): Not adequate to all contexts, vague, allowing errors in collected data due to different interpretations, inconsistent measures for aggregation across the organisation.		CI,CII, DS,G,B HP, K, FCB	MA SP1.3 MA SP1.4 MA SP2.1 MA SP2.2
	P11. First data collected were uncorrelated and too many indicators at the beginning.	R21: Conduct all necessary data collection cycles to find correlated metrics. R22: Give time for the metrics databases to become stable and allow the construction of relevant PPM and PPB. Different projects full cycles completed. R23: Categorise data. R24: Aggregate normalised data only for global view.	CIII, HP	OPP
	P12. Metrics Categorisation: Not all contexts data available. Unstable baselines without different categories. Not enough contextual information for data aggregation.		CI,CII, DS	OPP
	P13. Baselines not applicable to all projects		CI,CII	OPP SP1.3 OPP SP1.4 QPM SP2.2
Metrics Usage	*P14. Abusive elimination of outliers:* exceptional causes of variation occurring once per project or new baseline being established.	R25: Quarantine outliers which cause is not immediately identified. R26: Recognise data points that are not outliers but are unique and recurrent.	CI FCB	MA SP1.4 MA SP2.2
	P15. Not all projects are measurable: Not collecting data from projects with a data structure different from the standard. Not using all derived metrics because of lack of definition of base measures adequate to context.	R27: Base measures should be defined for different work and then normalised to allow calculating derived measures. R14	CI,CII, DS	MA SP1.3 MA SP2.3
	P16. Effort estimates: without using historical data of effort or size.	R28: Expert judgment is more adequate in certain circumstances. R29: Use any related historical data: size, phase duration, time spent on task. R30: Do iterative planning and use real time sampling of processes when there is no previous data available.	CII GJ, DS	PP SP1.2 PP SP1.4
	P17. People behaviour: inaccurate personal data reports. Resistance to collect new/additional data after ML3.	R12, R13, R31: Never use personal data to evaluate people.	CI,CII, DS	
Tools Setup	*P18. Tools setup:* Problems in tools after deployment. Using the tools in practice and in different projects contexts allowed to identify undetected problems and necessary improvements.	R32: Tools are improved when used in practice, save time for their setup. R33: When correcting tools defects that have impact in the metrics definition and the process, do not use the collected data to build PPB.	CI	OPM SP2.2 OPM SP2.3
	P19. Overhead: in tools usage (data collection not completely automatic). *P20. Tools requirements:* New needs still being identified, new tools still being developed.	R34: Once PPM and PPB are stable only collect data that is needed. R35: Use automatic and unperceived data collection systems, with limited human intervention (start/stop and confirm). R36: Guarantee that data collection is precise (discipline people and change their mentality).	CI HP, J	

Entry Conditions – When planning a move towards HML respect time to: have mature levels and institutionalised practices; understand and analyse the needs for HML; find correlations between variables; reach stable metrics, processes, tools and work habits; select meaningful performance indicators and gather enough stable data points to have statistically meaningful historical data. Organisations need to carefully plan business and process improvement objectives, temporal horizon and resources (time, internal and external human resources, tools, training, etc.). (R1) We used DS to show that having PPM experts available to work is related with understanding PPM and PPB results by managers using them.

In CIII the implementation plan was long and all activities needed to be executed on the estimated time. At first, they considered that if an activity had overrun the schedule, time could be recovered by shortening others. The thought was abandoned once they realised that time could not be shortened in any task.

HMLs only work with a stable base, hence, introducing their practices can only occur after ML 2 and 3 are mature and institutionalised (Leeson 2009) (R2). CI analysed gaps to address problems in lower maturity levels; however, those processes were affected by changes to implement HML and there should have been a new cycle for them to mature. In CIII the move from ML3 to ML5 was uninterrupted, so ML3 matured and did not erode in the meantime. Guarantee the involvement of an expert in quantitative methods (e.g., statistician), preferably with experience in software and if possible also in CMMI, who can help better understand processes behaviour and correlations between variables, along with providing adequate statistical tools to different contexts (R3). DS showed relations between PPM and PPB creators understanding CMMI and organisations achieving HML.

Introducing a Six Sigma initiative in the organisation eases the introduction of the statistical knowledge necessary to the organisation workers (R4). In CI, Six Sigma helped to gain insight of information needs to achieve quantitative goals, solve problems and design PPB and PPM.

There must be a top down and bottom up revision of the organisation's processes, improvements/innovations, goals and quantitative goals (R5). R3 to R5 were also followed by CI and CIII, as they were part of the CMMI implementation process.

Process Definition and Implementation – The implementation of the model should reflect the culture of the organisation, and not be a copy imposed on personnel. Processes definition should identify current processes (*as is*) and improvements (*to be*) so they reflect organisation's culture and people good practices (Leeson 2009) (R6). When defining processes it is important to involve the experts, including those who use the process to do their work: project, technical and quality managers; developers; testers, etc. (R7). R6 and R7 were followed by CI and CIII by first understanding the existent process, identifying gaps and involving internal experts and users in the definition of improvements and new processes.

In multicultural organisations and when acquiring new companies imposing processes can result in a loss of knowledge and resistance to change. Different business units should share practices used and lessons learnt. Each business unit would then gradually and naturally adopt the other's practices if they better fulfilled needs. This approach allows creating processes without losing good practices, benefiting from cultural differences (R8). CII applied R8 in a single direction, i.e., analysed other business units metrics in order to adopt the ones that could be applicable to their projects lifecycle.

There should be goals specific for different business units, departments and projects, which must be related to the organisation business goals (R9). Such setting allows having goals monitored at all levels, avoiding loss of visibility by middle management in each level (R10). R9 and R10 helped CI maintain the visibility of processes and projects at different organisation levels and were part of the CMMI implementation process in CIII.

Commitment from the entire organisation is essential, including top management, middle management and the people who are actually doing the work (Leeson 2009) (R11). Training needs to be adequate for each role and to include not only the *what to do*, *how to do* and *hands on* components but also the *why shall we do it*, *what will we achieve* and *how do we see it* (R12). Top management needs to set goals, plan gradual institutionalisation, monitor and reward (R14). For that it is essential that they understand the processes. To have people commitment it is crucial that they understand the new practices, which can be achieved by coaching projects and people (Humphrey 2006), guiding and accompanying them (R13). R11 to R13 were used by both CI and CIII. Regarding the training on benefits and how can they be seen we cannot be sure if it was effective. However, in CI the coach corrected people's mistakes instead of guiding them. R14 was used by CI and CIII, the dissemination of processes was gradual, as they were ready to be deployed directly from pilot projects to the entire organisation. However, when organisations are large they should consider even more gradual dissemination, spreading practices in a small group of projects and gradually involving new ones, which can be done also profiting from team members' mobility. DS showed that the incentives to people improving and working in MA were more frequent in organisations that achieved HML.

Metrics and processes definitions mature when used in practice because it is when problems arise that it becomes more evident how procedures can actually be done. It is necessary to give some time to let processes and metrics mature before producing their final versions. (R15) R15 was used by CIII.

Metrics Definition – To establish business objectives and identify the indicators of the processes performance, organisations can use methods such as the goal-driven measurement (Park, Goethert, and Florac 1996) (R16). R16 was used by both CI and CIII so there was a clear view of which metrics were used to monitor different levels of goals and what was their definition.

Understanding metrics is a process that is completed when projects are using them as the final processes define. It is utterly necessary to train the entire organisation without undervaluing the effort in such tasks (R18).

Measures need to be defined with a set of repeatable rules for collecting and unambiguously understanding data and what they represent (Florac, Carleton, and Barnard 2000), if different people use them differently, then their definition is inadequate. The level of detail of metrics needs to be completely defined, understood (R17) and to consider the different types of projects' context, including the technology used (R19). For example, in some technologies there are more KLOC, the time to execute unit tests is negligible, etc. Another example is project type: outsourced, maintenance and development projects, for instance, will have different measurement and control needs. Those factors affect the metrics definition. R17 was also followed by CI but definitions needed to mature to ensure unambiguous collection and interpretation. In CIII it was necessary to define new metrics for ML5 to have the desired confidence, because the integrity of existent data from ML3 could not be assured. With time the definition of metrics was improved to tune the process models.

Define basic software processes about which data should be collected, then concatenate and decompose data in different ways to provide adequate information at project and organisation levels (Kitchenham et al. 2006). If necessary, data should be normalised to make them visible to top management (R20, R24). It is preferable to begin with a sub-process executed often and with a small number of variables so results come faster (R21). When the process is stable, then extend to other processes and more complex ones (Florac, Carleton, and Barnard 2000). R21 was part of the CMMI implementation process of CIII.

Metrics databases take time to become stable and allow the construction of relevant PPM and PPB. The data need to be categorised (R23). (Florac, Carleton, and Barnard 2000) refer to this process as "separating or stratifying data that belong to the same cause system". Nonetheless, to have adequate categorisation it is necessary that the different projects fully cycle to completion. Either the organisation has a significant number of concurrent projects with small lifecycles or it begins to work with first limited baselines that evolve with time. (R22)

Pilot projects are useful for stabilising processes, procedures and tools. The way people use tools may change the way metrics should be collected. Only after those projects are over and the practices are clearly defined, will the organisation be ready for training, the processes/procedures and tools be fully and correctly documented and people be able to learn and apply the practices. Changes may then be deployed so that processes become institutionalised.

Metrics Usage – Certain outliers can be removed from databases but it is necessary to pay attention to those not immediately understood. They can indicate a process is having a new behaviour (better or worse performance), be a common situation or indicate the existence of a different process, with a different behaviour and therefore originate new sub-processes (Florac, Carleton, and Barnard 2000). One way of avoiding the error of abusively eliminating such outliers is to monitor the process without the outlier in parallel to the process with the outlier, then decide the most adequate action: *1)* perform CAR; *2)* eliminate the outlier; *3)* establish a new baseline because process performance improved; *4)* create new sub-processes, in case of having sub-processes. (Florac, Carleton, and Barnard 2000) give an example of how to do it. (R25, R26). DS showed a relation between achieving HML and checking data precision and accuracy.

Regarding effort estimation, expert judgment is more adequate in certain circumstances, in particular when there is absolutely no previous knowledge of the project (Grimstad and Jorgensen 2006) (R28). Effort estimation does not necessarily need to be based on KLOC to be based on historical data; it can be based on other size metrics, phase duration or the time spent on task (R27, R29); R29 was followed by CI. When no data are available at all do iterative planning, so that when data from previous cycle are available they can be used to plan the following (R30). CI followed R30 in Team Software Process^{SM} pilot projects, and didn't use personal data for evaluation purposes (R31). Even following R31, CI had difficulties to convince people to accurately report effort – that is why we suspect that showing the benefits on training may not have been effective.

Tools Setup – It is important to understand that tools need time to be set up, especially when evolving existent ones (R32). Regarding R32, on both CI and CIII the tools initially used were more rudimentary. As processes, metrics and performance models and baselines were defined more complex tools were adopted or implemented.

Manual data collection is time-consuming and error prone (Hamon and Pinette 2010), so it should be automated. To avoid overhead in the collection process, the information system needs to have limited human intervention, e.g., reporting effort and measuring code. Effort spent on different software applications for doing the tasks may be measured and part of the effort automatically labelled; the person only verifies and corrects eventual errors by the end of a block of tasks. This avoids forgetting to report effort or constantly interrupting tasks to manually report. The information system should be composed by automatic storage tools connected to the development environment (Johnson et al. 2005). (R35) In both CI and CIII R35 was followed but it is always difficulty to totally eliminate human intervention in effort report, especially when people have other tasks than just developing code, for example. It is imperative that data collection is precise, if it was not so previously, people need to change their mentality and display discipline (R36).

The data collected when correcting those tools defects which have impact on the definition of the metrics and of the process should not be used to build PPB, because the process is not stable. (R33, R34) Regarding R34 the experience in CIII was that initially it was necessary to collect data of all variables they felt could be important to create models and establish baselines. In time the non-used metrics were abandoned, leaving only the necessary ones.

In a follow up of CII we found that R18 and R27 were used when they defined their specific metrics and implemented the estimation tool.

Validity and Limits to Generalisation

Due to access limitations, the three case studies had a different design so they cannot be considered multiple-case studies (Yin 2009). Only part of the design of CI was repeated on CII, and in CIII we only interviewed a consultant involved in the appraisal. We can classify it has a semi-multiple case study. In CI and CII we used multiple sources of evidence, assuring construct validity. However in CIII we could not assure it. In all cases we had our results reviewed by key informants. To ensure internal validity we did pattern matching by classifying information and aggregating it under each category; built explanations and addressed rival explanations. External validity was partially tested by replicating part of the design used in CI in CII, and analysing DS. Nonetheless, for each case study we used theory.

Regarding limits to generalisation, we only analysed three cases but some of the problems that we identified were also found in the literature and DS. Subsequently, we consider that these problems can be common to other organisations implementing CMMI, measurement programs or doing software process improvements.

Conclusions

Concerning CMMI problems, the DoD expressed the necessity to "Develop meaningful measures of process capability based not on a maturity level, e.g. Level 3, but on process performance" (Schaeffer 2004). CMMI V1.3 is more focused on performance of organisations but SCAMPI is becoming more efficient (Philips 2010), as it reduced the number of necessary evidence – eventually increasing the probability of leaving problems undetected.

The difficulties in implementing CMMI, in particular HML, are common to the problems found on metrics programmes and software process improvements in general. In particular, Software Engineering metrics are still ambiguous (Goulão 2008; Breuker et al. 2009), impeding an implementation common to all organisations. With the objective of understanding CMMI problems better, we did the research presented on this paper and compiled them. We also assembled a set of recommendations useful for software development organisations.

Interestingly enough, part of the identified problems is rooted in the CMMI lower maturity levels (2 and 3) as they must be stable before moving to high maturity. The evidence we analysed show that the problems were common to the different sources of information. In the case of the ones we found in the case studies, we verified that they also occurred in other organisations. Furthermore, there are also several problems in the Measurement and Analysis process area that become more evident when implementing CMMI maturity level 4, as they affect process performance models and baselines.

There is a wide variety of methods to implement CMMI practices. As the model is just a guide which tells you *what to do*, but not *how to do it*, room is left for various implementations that may not always lead to the desired performance results. Moreover, SCAMPI's objectives do not include appraising performance. Consequently, problems and difficulties can occur when implementing CMMI, some of which can persist after appraisal. As future research we are providing a framework for self-assessing the quality of implementation of CMMI practices and effects of improvements, based on compliance, efficiency and effectiveness (Lopes Margarido et al. 2012). We believe our framework help prevent implementation problems and allows better control of organisations performance.

Acknowledgements

This research was partially sponsored by Fundação para a Ciência e a Tecnologia (FCT): Programa Operacional Potencial Humano (POPH) of QREN, and Fundo Social Europeu (FSE).

We would like to thank the organisations that received us, all our interviewees, the consultants that discussed this topic with us and all our reviewers for their contribution to this work.

References

Armstrong, James, Richard Barbour, Richard Hefner, and David H. Kitson. 2002. Standard CMMI^{SM} Appraisal Method for Process Improvement (SCAMPI^{SM}): Improvements and Integration. *Systems Engineering* 5 (1):19-26.

Barcellos, Monalessa Perini. 2009. Uma Estratégia para Medição de Software e Avaliação de Bases de Medidas para Controlo Estatístico de Processos de Software em Organizações de Alta Maturidade. Doctoral, Engenharia de Sistemas e Computação, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

Breuker, Dennis, Jacob Brunekreef, Jan Derriks, and Ahmed Nait Aicha. 2009. Reliability of software metrics tools. In *International Conference on Software Process and Product Measurement*. Amsterdam.

Charette, Robert, Laura M. Dwinnell, and John McGarry. 2004. Understanding the Roots of Process Performance Failure. *CROSSTALK The Journal of Defense Software Engineering* 17 (8):18-22.

Chrissis, Mary Beth, Mike Konrad, and Sandy Shrum. 2011. *CMMI for Development®: Guidelines for Process Integration and Product Improvement*. 3 ed, *SEI Series in Software Engineering*. Massachusetts: Addison-Wesley.

CMU/SEI. *Criteria for Audits of High Maturity Appraisals*. CMU/SEI 2008 [cited 21-06-2012. Available from http://www.sei.cmu.edu/cmmi/solutions/appraisals/himataudits.cfm.

CMU/SEI. 2011. Standard CMMI® Appraisal Method for Process Improvement (SCAMPI^{SM}) A, Version 1.3: Method Definition Document. CMU/SEI.

Florac, William A., Anita D. Carleton, and Julie R. Barnard. 2000. Statistical Process Control: Analyzing a Space Shuttle Onboard Software Process. *IEEE Softw.* 17 (4):97-106.

Goldenson, Dennis R., Diane L. Gibson, and Robert W. Ferguson. 2004. Why Make the Switch? Evidence about the Benefits of CMMI. In *SEPG*: CMU/SEI.

Goldenson, Dennis R., James McCurley, and Robert W. Stoddard II. 2008. Use and Organizational Effects of Measurement and Analysis in High Maturity Organizations: Results from the 2008 SEI State of Measurement and Analysis Practice Surveys. CMU/SEI.

Goulão, Miguel Carlos Pacheco Afonso. 2008. Component-Based Software Engineering: a Quantitative Approach. Doctoral, Departamento de Informática, Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia, Lisboa.

Grimstad, Stein, and Magne Jorgensen. 2006. A framework for the analysis of software cost estimation accuracy. In *Proceedings of the 2006 ACM/IEEE international symposium on Empirical software engineering*. Rio de Janeiro, Brazil: ACM.

Hamon, Patrick, and Olivier Pinette. 2010. Les indicateurs Mesure & Analyse. Spirula.

Hollenbach, Craig, and Doug Smith. 2002. A portrait of a CMMI^{SM} level 4 effort *Systems Engineering*, 52-61.

Humphrey, Watts S. 2006. *TSP^{SM}: Coaching Development Teams*, *The SEI Series in Software Engineering*: Addison-Wesley.

Johnson, Philip M., Hongbing Kou, Michael Paulding, Qin Zhang, Aaron Kagawa, and Takuya Yamashita. 2005. Improving Software Development Management through Software Project Telemetry. *IEEE Softw.* 22 (4):76-85.

Kitchenham, Barbara, Cat Kutay, Ross Jeffrey, and Colin Connaughton. 2006. Lessons Learnt from the Analysis of Large-scale Corporate Databases. In *International Conference on Software Engineering*. Shanghai: ACM.

Leeson, Peter. 2009. Why the CMMI® does not work. In *SEPG Europe*. Prague, Czech Republic: CMU/SEI.

Lopes Margarido, Isabel, João Pascoal Faria, Raul Moreira Vidal, and Marco Vieira. 2012. *Towards a Framework to Evaluate and Improve the Quality of Implementation of CMMI® Practices*. Vol. 7343, *Product-Focused Software Development and Process Improvement*. Madrid: Springer Berlin / Heidelberg.

Masters, Steve, PhD Sandi Behrens, Judah Mogilensky, and Charlie Ryan. 2007. SCAMPI Lead AppraiserSM Body of Knowledge (SLA BOK). In *Software Engineering Process Management Program*: CMU/SEI.

McCurley, James, and Dennis R. Goldenson. 2010. Performance Effects of Measurement and Analysis: Perspectives from CMMI High Maturity Organizations and Appraisers CMU/SEI.

Monteiro, Luis Felipe Salin, and Kathia Marçal de Oliveira. 2011. Defining a catalog of indicators to support process performance analysis. *Journal of Software Maintenance and Evolution: Research and Practice* 23 (6):395-422.

Park, Robert E., Wolfhart B. Goethert, and William A. Florac. 1996. Goal-Driven Software Measurement - A Guidebook. CMU/SEI.

Philips, Mike. 2010. CMMI V1.3 Planned Improvements. In *SEPG Europe 2010*. Porto, Portugal: CMU/SEI.

Pyster, Arthur. 2006. What Beyond CMMI Is Needed to Help Assure Program and Project Success?

Unifying the Software Process Spectrum, edited by M. Li, B. Boehm and L. Osterweil: Springer Berlin / Heidelberg.

Radice, Ron. 2000. Statistical Process Control in Level 4 and Level 5 Software Organizations Worldwide. In *Software Technology Conference*: CMU/SEI. Original edition, Proceedings of the 12th SEPG Conference, Seattle, Washington, March, 22, 2000.

Schaeffer, Mark. 2004. DoD Systems Engineering and CMMI. In *CMMI Technology Conference and User Group*.

Stoddard II, Robert W., and Dennis R. Goldenson. 2009. Approaches to Process Performance Modeling: A Summary from the SEI Series of Workshops on CMMI High Maturity Measurement and Analysis. CMU/SEI.

Takara, Adriano, Aletéia Xavier Bettin, and Carlos Miguel Tobar Toledo. 2007. Problems and Pitfalls in a CMMI level 3 to level 4 Migration Process. In *Sixth International Conference on the Quality of Information and Communications Technology*.

Yin, Robert K. 2009. *Case Study Research Design and Methods*. Fourth ed, *Applied Social Research Methods Series*: SAGE.