International Conference on Software Process and Product Measurement (MENSURA'96), November 6-8, 2006, Cádiz, Spain, Edited by Alain Abran, Reiner Dumke and Mercedes Ruiz, pp.252-263, Publicaciones de la Universidad de Cádiz, 2006. (ISBN13: 978-84-9828-101-9, ISBN10: 84-9828-101-6).

Towards an Internal Numerical Taxonomy of Software Process Assessment Methods

Alberto Sampaio[1], Edwin Gray[2], Mário Martins[3], Isabel B. Sampaio[4]

[1] Departamento de Engenharia Informática and
Laboris (Research Laboratory on Systems)
Instituto Superior de Engenharia do Porto, 4200-072 Porto, Portugal.
as@dei.isep.ipp.pt
www.{laboris, dei}.isep.ipp.pt
[2] School of Computing and Mathematical Sciences
Glasgow Caledonian University, Cowcaddens Road,GLASGOW G4 0BA, UK.
e.gray@gcal.ac.uk
www.gcal.ac.uk
[3] Departamento de Informática
Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal.
fmm@di.uminho.pt
www.di.uminho.pt
[4] Departamento de Engenharia Informática
Instituto Superior de Engenharia do Porto, 4200-072 Porto, Portugal.
is@dei.isep.ipp.pt
www.dei.isep.ipp.pt

Abstract. *Objectives*: The main goal of this paper is the internal comparison and classification of software process assessment methods. *Methods*: This exploratory study used a new methodology, based on numerical taxonomy. Eight software process assessment methods (SCAMPI; RAPID; SPICE; EVALUATION; BOOTSTRAP; SA-SI; MMA, and CBA-IPI) were classified. For the study were used documents describing the assessment methods. *Results*: A data table with a final set of 112 characters; proximity matrices of methods from step 4, and the final result was a hierarchical taxonomy of the eight methods selected from a set of taxonomies (step 5) validated (step 6) for a significance level of 5%. Conclusions: The internals comparison and classification produced are preliminary and are in accordance with our perceptions about the internal relationships between the methods. From the taxonomy was concluded that the methods SCAMPI and Bootstrap are the closest, followed by CBA-IPI, RAPID and MMA, Evaluation, SA-SI and SPICE by this order. Because of the exploratory nature of the work, the need for more studies is pointed.

1 Introduction

Software process assessment (SPA) is an important step in the software process improvement (SPI) cycle. Because SPA results are the input for the SPI initiatives, a "good" assessment can be seen as one that allows a successful improvement. To know the reasons why an assessment method contributes to such a "good" assessment it is necessary to know which attributes, and their combinations, that successful methods possess and that the others do not. So, it is necessary to compare current methods about their internal characteristics, as done in this paper. Such comparisons, together with empirical data about the use of the methods, will allow the discovery of which characteristics should be present in "good" SPA methods.

Software process improvement, in general, and the assessment area, in particular, are far from being guided by a scientific theory and could not be considered scientifically mature. In a comparison or classification of SPA methods, if it is lacking a theoretical base, then it is not rational to prefer some characteristics in detriment of others. This implies the need for a comparison and classification method where all known characteristics could be considered.

Some comparisons and classifications have been made about strategies and models of process improvement, but very few of SPA methods are described in the literature. Additionally there are no well defined methods for comparing SPA methods, as there is a lack of approaches for software methods in general [39]. Those methods belong, in general, to the second and fourth classes of Henk Sol [38] types of methods of comparison. The goal of our work is the comparison and classification of SPA methods based on their internal characteristics. The comparison and classification was done with a new methodology based on numerical taxonomy, as proposed initially for biological sciences, which solves the previous limitations.

In the next sections we present related work, the research approach and its application, after which we discuss results and present some conclusions.

2. Background and Related Work

It is possible to find a reasonable number of studies comparing SPA models, as [2], [13], [14], [22], [24], [25], [26], [33], [41], [42], [43] [44], [47], but only a reduced number of comparisons of SPA methods. The comparisons of SPA methods we could find reported in the literature are, Ares et al. [1], Haase [12], Rout[28], Rout and Gasston [29], Rout and Nielson [30], Wang et al. [45], Zahran [47], and the Software Engineering Institute (SEI) [36].

However there are no simultaneous studies of comparison and classification and none of the studies could be considered adequate: usually the studies analyze only a few methods or a few characteristics depending of each author's criteria or goals and without a well defined comparison process, with all the consequent problems, like lack of uniformity, difficult to determine its quality, difficult of construction and impossibility to replicate the classification. Replication studies are considered important in an empirical software engineering but, usually, are not developed [3].

It was necessary to use a new approach for the comparison and classification of software methods, and SPA methods in particular. The new methodology used allows the creation of a numerical taxonomy of the methods. Because it analyzes methods, the approach is referred as a methodology. The approach is considered empirical, statistical and multidimensional. It is based on a numerical taxonomy used in biology that was, mainly proposed in the end of the fifty's by Sokal and Sneath [37] to classify live beings. Despite its origin, the numerical taxonomic methods are considered applicable to other areas [20] [15].

The methodology advocates the importance of gathering as many characteristics as possible about each entity (method) to be classified in order to produce a “natural” classification. This is known as the saturation principle and that is consistent with the lack of a theoretical base as stated previously. Because it is a rigorous and numerical approach, studies that use it could be more easily replicated.

3. The Comparison and Classification Methodology

In the methodology the resulting classification is hierarchical. Each entity is designated by OTU (“Operational Taxonomic Unit”) and the characteristics by characters. The methodology comprises the following steps:

1. Choose the OTU’s – In this step the SPA are chosen.
2. Character discovery and measurement - Consists of determining which characters should be included and the “measurement" for each OTU.
3. Codification – Codification of the characters and their normalization; It also involves the treatment of missing data.
4. Proximity – In this step the proximity matrix is produced showing the dissimilarities between the OTU’s. This gives a value of “global affinity” between the methods. In the calculation several coefficients of dissimilarity are used.
5. Clustering – This step try to group the methods. Several clustering criteria are applied in a hierarchical agglomerative strategy. This is the cluster analysis step.
6. Validation - Statistical validation of the results in order to find the best solution. Validation is done based on the idea of stability.
7. Interpretation – Interpretation of the solution – It includes the formation of potential groups and conceptualization. This step requires knowledge of the application domain of the study (SPA methods in the present case).
8. a) Data extraction - given a taxonomic group, obtain the data about that group; b) Identification of cases - determination of the respective group of a given method.

The first three steps allow the creation of a data table with the characters for each method, which is the usual result of comparative studies. With this methodology we can go a step further by allowing the development of a classification of the entities. With the methodology interpretation is possible at the various steps, but it is done mainly at the seventh. The authors of the methodology advocate the application of several measures at step 4, and of several methods (criteria) at step 5, in order to produce several classifications. The consistency between the several results is considered an indicator of the quality of the solution. The classification with the smallest error could be selected as the solution.

4. The Study

The study is exploratory and followed the methodology steps, rules and recommendations. These rules and recommendations affected the decisions. Step number seven is not covered because it was not completely developed and because it consists mainly of numbers identifying the characters belonging to each of the classes found making no sense without the data table. The last step is not part of the study.

Given the high quantity of numerical computation required, the Matlab software package was used. Notwithstanding the enormous functionality offered by the package, notably from its "statistics toolbox", it was necessary to develop many new Matlab functions using its built-in language. Routines were developed for numeric processing at steps, 3, 4, 5, 6 and 7. Computation was done in a 1GHz laptop with 2*256 MB of main memory and Microsoft Windows XP operating system.

4.1. Choose the OTU's (Step 1)

The study attempted to include all the SPA methods that have some documentation describing them. However it was not possible to include all the SPA methods, because there was not enough information available for some of them. In some cases the methods were not freely available. Older versions of existing methods have not been included. Because methods using only a questionnaire are considered too limited [11] they have not been included. There was also a set of methods planned to be included later in the classification (step 8) and that will allow test classification sensibility to new cases. Eight methods were chosen. From a sampling theory point of view it is an intentional non-probabilistic sample. The methods selected and their main reference sources are listed in Table 1. The set of methods has dimension n=8.

Table 2. Methods included and their sources.

Method	Author	Main sources
1. SCAMPI	SEI	[35],[36]
2. RAPID	T. Rout et al.	[31]
3. 15504	ISO/IEC	[10],[19]
4. Evaluation	J. Ares et al.	[1]
5. Bootstrap 3	Bootstrap Institute	[4],[5],[40]
6. SA-SI	M. Hobday, T. Brady	[17]
7. MMA	Karl Wiegers, D.Sturzenberger	[46]
8. CBA-IPI	SEI	[9],[21]

4.2. Character Discovery and Measurement (Step 2)

Character discovery was done OTU by OTU, by reviewing exhaustively all the documents related to the methods in order to find items that could be recognized as characteristics of a method. The discovery began with the SCAMPI method, because it was one of the biggest, this is, more documented. This was easy because SCAMPI requirements were all well stated in a document [36] from SEI. Our list of characters reflects that. The set of characters about the ISO/IEC 15504, was obtained through a table [36], connecting the SCAMPI method with 15504, and with the ISO/IEC15504 documentation. Because character discovery is empirical, each character being discovered by observation of its method there are no empty variables. A total of 165 characters were discovered.

For the cases where the character presents some well-defined values, these values were used. If not, a dichotomous classification scale, like “Present” and “Absent” was adopted by default. Some times the two state scale gave place to an ordinal scale of the kind, “Absent”, “Partially Present” and “Present”. The scale in these cases is ordinal. In other situations there was the need to use nominal multi-state and quantitative data. With the discovery and measurement of the characters a data table[21] with the 165 characters was constructed. An extract of three characters of the data table is presented in Table 2, where “S” means “Present” and “N” “Absent”.

Table 3. An example of three characters.

Nº	Characters \ Methods	1	2	3	4	5	6	7	8
1	The Sponsor is a Senior manager	S	S	S	N	S	S	N	S
2	The Sponsor is a manager, but not Senior	N	N	N	S	N	N	N	N
3	The Sponsor is from the SPEG	N	N	N	N	N	N	S	N

4.3. Data Codification (Step 3)

The data were codified with the integer values {0,1} for the {S,N} case, to {0,1,2} for the {S,P,N} case, and no change was made for quantitative values. Next it was standardized in amplitude to the interval 0 to 1. There was a significant number of missing data and it did not make sense to apply a statistical procedure to treat the missing data. Note that there is about 1/20 ratio of entities by variable, but regression methods usually require much more cases than variables. So, it was applied the simplest solution, variable-deletewise, that means to delete the variables with missing data. After the deletion a total of p=112 characters remained.

4.4 Proximity (Step 4)

In this step the dissimilarity matrices were calculated by applying 5 coefficients (Euclidean, Standardized Euclidean, City Block, Minkowski with p=0.5 and Minkowski with p=4). These distances assume that data is quantitative. The resulting proximity matrices are shown in the Appendix. Each entry represents a distance between a pair of methods, in a total of $N = n\ (n\text{-}1)\ /\ 2 = 28$ distances.

---

[21] The table, in Portuguese, is available from the first author.

4.5 Clustering (Step 5)

For clustering, four hierarchical criteria (Single Linkage; Complete Linkage; Average; and Ward) were applied. These criteria could produce very different classifications and induce ultrametrics. The result was a total of twenty classifications. The classifications are presented in the form of indexed trees, called dendrograms. The index is a function describing the strength of the connection between nodes.

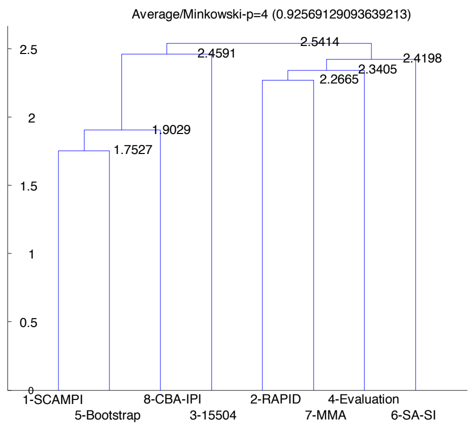

Fig. 3. The dendrogram of classification number 19.

4.6. Validation (Step 6)

The goal is to find the classification with the small error introduced during the classification, defined by the distortion introduced by the clustering method. The statistic chosen was the coefficient of cophenetic correlation (CCC) which is complementary to the clustering error. In Fig.1 the value between parentheses is the CCC. The significance of the results obtained was tested for each method for a significance value of 0.05. Because the CCC distribution depends from application to application, the null distribution for each method was determined by Monte Carlo sampling, with a number of 1000 samples. From the twenty initial classifications, five were rejected.

The concordance between the dendrograms of the statistical significant classifications was analyzed by applying the Kendall's W coefficient of concordance to their topologies (described by the Partition Membership Divergence [27] topological descriptor). The value obtained confirmed concordance for a significance level of 0.01. To select the classification with the smallest error it was necessary to select the classification with the highest CCC. Such solution is the number 19, Average/Minkowski-p=4, which is showed in Figure 1. But the solution number 3, Average/Euclid, has nearly the same value (equal to the third decimal place).

5. Discussion and Validity of Results

Methods SCAMPI and Bootstrap from the solution were the first to be joined, followed by CBA-IPI. Next RAPID and MMA were joined, followed by Evaluation and SA-S.I In the end SPICE joined the first group. The proximity of Bootstrap and SCAMPI is understandable because both methods were developed from a previous SEI method and from the SEI CMM model. They are also ISO/IEC 15504 compliant. CBA-IPI is also a SEI method, and SCAMPI was directly based on it. Methods RAPID and MMA are the most light. SA-SI is more problem oriented and depends only partially on an assessment model (the CMM). The Evaluation method is the only one based on a theory (theory of evaluation). SPICE is a standard and so it possess the minimum set of requirements that methods should have to be SPICE compliant. This could explain the distance of SPICE.

Some results obtained from the matrices followed by possible explanations:

1. In the dissimilarity matrices the most similar methods are 1 and 5, SCAMPI and Bootstrap, or 1 and 8, SCAMPI and CBA-IPI. Between the more dissimilar methods the variability is higher.
2. In the matrices, CBA-IPI appears closer of SCAMPI than of Bootstrap.
3. SPICE appears nearly equidistant of SCAMPI and Bootstrap.
4. The methods MMA e RAPID appears always near each other.

Possible explanations, in the same order:

1. These proximities were expected as explained earlier.
2. Both SACMPI and CBA-IPI were developed by SEI, but Bootstrap was not.
3. Bootstrap is compliant with 15504, but much more complete. SCAMPI uses CMMI models but it can be compliant with the standard if used with the continuous version of the CMMI.
4. Again, these two methods are less heavy than the others. However RAPID is SPICE based but MMA is CMM oriented. MMA is also more flexible.

Some validity issues should be pointed. The main limitations of the study are of two kinds: limitations with the numerical taxonomy itself; and construction of the data table by one person only, the first author. Character discovery is a subjective task based on the analyst intuition and knowledge of the domain. However, was possibly compensated by the large number of characters as advocated by the methodology. Because character discovery was done by only one person we couldn't evaluate the validity of that process. The differences in the amount and quality of information between the methods documentation could also have affected the results. The character discovery and measurement started with the SCAMPI method described in the document ARC-Assessment Requirements for CMMI [36] and its use probably introduced some bias in the character discover process. Because several requirements have been divided, some deleted and new ones added, hopefully we reduced a possible bias. Finally, missing data was in a significant number. All this implied that the results could not be considered definitive.

3. Conclusions and Future Work

The goal of this study was the development of a comparison and classification of SPA methods based on their internal characteristics. It was used an innovative approach based in the numerical taxonomy as used in biology to classify live beings. An internal classification is in agreement with measurement theory because external measurement depends on the measurement of internal attributes. Direct measures occur in an initial phase of scientific development [48], as it is the case with SPA. As long as we know, this is the first numerical comparison and classification of SPA methods, or of software methods of any kind based on their internal characteristics.

The classificatory study was two-fold exploratory. From one side some SPA methods have never been compared and some characteristics never included, from other side, it was the first time this new comparison and classification methodology has been used. In the study eight SPA methods were compared and classified. The classification was done based on 112 characters. The chosen solution was the hierarchical taxonomy number 19. Potential groups were not created because of the exploratory nature of the study. However, it was apparent a group of three methods, SCAMPI, Bootstrap and CBA-IPI.

On this stage of work we were not motivated by practical reasons, however, with the actual results is possible to choose a method in a specific situation by being aware of the relevant characters and their importance. With this information the user could consult the data table to know whose methods better satisfy such characters. Or, using information from step 7, he could inspect in the taxonomy the classes that possess the relevant characters.

A future review of the SPA methods could allow a reduction of missing data, and then, another strategy for missing data should be considered. The subjective task of character discovery also needs to be improved. The binary and ordinal scales were treated as interval scales, because such violations could be considered acceptable in an exploratory study [6] and we did not find that the effects of such violations did produce significant differences in the results. Obtaining more documentation for the study proved to be difficult.

This work is included in a more wide work intended to identify the relationships between the internal characters and the results produced by the methods (the external characteristics). This means to apply the methodology, or part of it, two more times, to detect external characters and to relate external with internal characters. So other studies should be developed. To use the methodology in such staged way means that it can be seen as an empirical research methodology that could allow detect cause-effect relations. About the methodology it is planned to complete step 7 with concept formation and the study of the validity of the saturation principle with software methods documentation. We also intend to apply the methodology to other kinds of software methods and software phenomena.

Acknowledgement

The authors would like to thank Barbara Kitchenham for the remarkable comments to a previous version of this paper contributing in a significant way for its improvement.

References

[1] Ares, J., Garcia, R., Juristo, N., López, M., Moreno, A, A More Rigorous and Comprehensive Approach to Software Process Assessment, Software Process: Improvement and Practice, John Wiley & Sons, Vol.5, I.1, 3-30, 2000.

[2] Bamford, R.C., Deibler II, W.J., Comparing, Contrasting ISO 9001 and the SEI Capability Model, IEEE Computer, Vol.26, N., 10, pp.68-70, October, 1993.

[3] Basili, V., Shul, F., Lanubile, F., Building Knowledge through Families of Experiments, IEEE Transactions on Software Engineering, pp.456-473, Vol.25, N.4, July/August, 1999.

[4] Bicego, A., Khurana, M., Kuvaja, P., Bootstrap 3.0 Software Process Assessment Methodology, In C. Hawkins , M. Ross , G. Staples (Editors), Software Quality Management VI: Quality Improvement Issues, Springer-Verlag, NY, 1998.

[5] Bootstrap, Bootstrap: Europ's Assessment Method, Software, Vol.10, N.3, May, 1993.

[6] Briand, L., El-Emam, K., Morasca, S., On the Application of Measurement Theory in Software Engineering, Empirical Software Engineering: An intl. Journal, Vol.1, N.1, 1996.

[7] CSTB - Computer Science and Technology Board (National Research Board from the National Academy of Sciences), Scaling Up: A Research Agenda for Software Engineering". CACM, Vol.33, N.3, pp.281-293, 190.

[8] Dion, Raymond, "Process Improvement and the Corporate Balance Sheet," IEEE Software, July 1993, pp. 28-35.

[9] Dunaway, D., Masters, S., CMM-Based Appraisal for Internal Process Improvement (CBA IPI) V1.1: Method Description, Software Eng. Institute, CMU/SEI-96-TR-007, 1996.

[10] El Amam, K., Drouin, J., Melo, W., (Editors), SPICE: The Theory and Practice of Software Process Improvement and Capability Determination", IEEE CS Press, 1998.

[11] Gray, E.M., Sampaio, A., Benediktsson, O., An Incremental Approach to Software Process Assessment and Improvement, Proceedings of the British Computer Society Quality Special Interest Group's 11th Annual International Software Quality Management Conference: 'SQM 2003', 23rd - 25th of April, 2003, Glasgow Caledonian University, Glasgow, Scotland, 2003.

[12] Haase, V. H., Software process assessment concepts. Journal of Systems Architecture, the EUROMICRO Journal, Vol.42, N. 9, pp 621-631, Dec., 1996.

[13] Hailey, V.A., A Comparison of ISO9001 and the SPICE Framework, In K. El Amam, J. Drouin, W. Melo, (Editors), SPICE: The Theory and Practice of Software Process Improvement and Capability Determination", IEEE CS Press, 1998.

[14] Halvorsen, C.P., Conradi, R., A Taxonomy of SPI Frameworks, Twenty-Fourth Annual Software Engineering Workshop, SEL/NASA/GSFC, 1-2 December 1999.

[15] Hartigan, J.A., Classification, In Kotz, S., Johnson, N.L. (Editors-in-Chief), Encyclopedia of Statistical Sciences, (pp.1-10), Vol.2, John Wiley & Sons, Inc., 1982.

[16] Hetzel, B., The Sorry State of Software Practice Measurement and Evaluation, in Fenton, N., Whitty, R., Iizuka, Y. (Eds.), Software Quality Assurance and Measurement: A Worldwide Perspective, Thomson Comp. Press, 1995.

[17] Hobday, M., Brady, T., A Fast Method for Analysing and Improving Complex Software Processes, R & D Management, Vol.30, N.1, (pp.1-21), Blackwel Publishers Ltd, 2000.

[18] Humphrey, W.S., T.R. Snyder, and R.R. Willis, "Software Process Improvement at Hughes Aircraft," IEEE Software, July 1991, pp. 11-23.

[19] ISO/IEC TR 15504 - Information technology - Software process assessment — Parts 3,4,5, 1998.

[20] Jardine,N., Sibson, R., Mathematical Taxonomy, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1971.

[21] Masters, S., Bothwell, C., CMM Appraisal Framework, Version 1.0, Software Engineering Institute, CMU/SEI-95-TR-001, 1995.

[22] Messnarz, R., A Comparison of BOOTSTRAP and SPICE, IEEE Software Process Newsletter, N.8, Winter, 1997.

[23] Osterweil, L.J., Software Processes Are Software Too, Revisited: An Invited Talk on the Most Influencial Paper of ICSE 9, pp.540-548, Proceedings of the 19th International Conference on Software Engineering, May 17-23, 1997, Boston, USA, ACM, 1997.

[24] Paulk, M.C., Comparison of ISO 9001 and the Capability Maturity Model for Software, Software Engineering Institute, CMU/SEI-94-TR-012, 1994.

[25] Paulk, M.C., Analyzing the Conceptual Relationship Between ISO/IEC 15504 (Software Process Assessment) and the Capability Maturity Model for Software, 1999 International Conference on Software Quality Cambridge, MA, 1999.

[26] Paulk, M.C., Konrad, M.D., Garcia, S.M., "CMM Versus SPICE Architectures," IEEE Software Process Newsletter, No. 3, Spring, 1995a.

[27] Podani, J., Dickinson, T.A., Comparison of Dendrograms: a multivariate approach, Canadian Journal of Botany, Vol.62, pp.2765-2778, 1984.

[28] Rout, T.P., SPICE and the CMM: Is the CMM Compatible with ISO/IEC 15504?, AquIS '98, Venice, Italy, March, 1998.

[29] Rout, T.P., Gasston, J.L., Different Approaches to Software Assessment, In T.P. Rout (Ed.), Software Process Assessment and Improvement, Computational Mechanics Pub., 1998.

[30] Rout, T.P., Neilson, M.P., Gasston, J.L., Experiences with the use of different approaches to software process assessment, Proceedings of the 2nd International Conference on Software Quality Management. Edinburgh. Scotland, 1994.

[31] Rout, T.P., Tuffley, A., Cahill, B., Hodgen, B., The Rapid Assessment of Software Process Capability, Proceedings of SPICE 2000 Intl. Conference, Limerick, Ireland, 2000.

[32] Rozum, J., Concepts on Measuring the Benefits of Software Process Improvements, Software Engineering Institute, CMU/SEI-93-TR-009, 1993.

[33] Saiedian e McClanahan, Frameworks for Quality Software Process: SEI Capability Maturity Model versus ISO 9000, Software Quality Journal, Vol.5, pp.1-23, 1996.

[34] Sampaio, A., Gray, E.M., Martins, M., A Comparison of SPA Methods, 31ts EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA), WiP session , August 31st-September 3rd, 2005, Porto, Portugal, 2005.

[35] SEI (CMMI Product Development Team), SCAMPI(SM), V1.0 Standard CMMI SM Assessment Method for Process Improvement: Method Description, Version 1.0, Technical Report CMU/SEI-2000-TR-009, 2000.

[36] SEI (CMMI Product Development Team), ARC, V1.0 Assessment Requirements for CMMI, Version 1.0, Technical Report CMU/SEI-2000-TR-011, 2000.

[37] Sokal, R.R., Sneath, P.H.A, Principles of Numerical Taxonomy, W. H. Freeman and Company, 1963.

[38] Sol, H.G., A Feature Analysis of Information Systems Design Methodologies: Methodological Considerations, pp.1-7, in T.W. Olle, H.G. Sol, C. Tully (Editors), Information Systems Design Methodologies: A Feature Analysis, IFIP WG8.1 Working Conference, 5-7-July, 1983, UK, North-Holland, 1983.

[39] Song, X., Osterweil, L.J., Toward Objective, Systematic Design-Method Comparisons, IEEE Software, Vol. 9, N.3, pp.43-53, May, 1992.

[40] Stienen, H., Engelman, F., Lebsanft, BOOTSTRAP: Five Years of Assessment Experience, In Proceedings of the 8th Int. Workshop on Software Technology and Engineering Practice (STEP'97), IEEE Press, 1997.

[41] Tingey, M.O., Comparing ISO 9000, Malcolm Baldridge, and the SEI CMM for Software: A Reference and Selection Guide, Upper Saddle River: Prentice-Hall, 1997.

[42] Tully, C., Kuvaja, P., Messnarz, R., Software Process Analysis and Improvement: a Catalogue and Comparison of Models, In Richard Messnarz, Colin Tully (Eds), Better Software Practice for Business Benefit, IEEE Computer Society, Los Alamitos, California, 1999.

[43] Wang, Y., Court, I., Ross, M., Staples, G., King, G., Dorling, A., Quantitative Analysis of Compatibility and Correlation of the Current SPA/SPI Models, In Proceedings of the 3rd Int. Software Engineering Stantards Symposium (ISESS'97), IEEE Press, 1997a.

[44] Wang, Y., Court, I., Ross, M., Staples, G., King, G., Dorling, A., Quantitative Evaluation of the SPICE, CMM, ISO 9000 and BOOTSTRAP, In Proceedings of the 3rd Int. Software Engineering Stantards Symposium (ISESS'97), IEEE Press, 1997b.

[45] Wang, Y., King, G., Dorling, A., Wickberg, H., King, G., Experience in Comparative Process Assessment with Multi-Process-Models, Proceedings of the 25th Euromicro Conference (EUROMICRO '99), IEEE Press, 1999c.

[46] Wiegers, K.E., Sturzenberger, D.C., A Modular Software Process Mini-Assessment Method, IEEE Software, Vol.17, N.1, 2000, pp.62-69.

[47] Zahran, S., Software Process Improvement: Practical Guidelines for Business Success, Addison-wesley, 1998.

[48] Zuse, H. A Framework of Software Measurement, Walter de Gruyter&C., Berlin, 1997.

Appendix: Proximity Matrices

Euclidean

6.94	5.79	6.38	3.28	7.09	7.06	3.28
	5.71	5.82	6.55	6.18	5.50	6.92
		5.89	5.50	6.69	7.03	6.38
			6.20	6.04	5.71	5.96
				6.89	6.83	4.30
					6.29	6.71
						6.64

CityBlock

51.34	36.00	42.50	12.50	54.50	55.83	11.50
	36.34	36.84	44.84	43.16	35.49	50.84
		37.50	31.50	50.50	54.83	43.50
			41.00	41.00	38.33	38.00
				50.00	53.33	21.00
					45.33	49.00
						49.33

Minkowski p=4

2.60	2.37	2.51	1.75	2.62	2.60	1.79
	2.34	2.37	2.68	2.42	2.27	2.60
		2.39	2.51	2.52	2.59	2.49
			2.46	2.40	2.31	2.41
				2.74	2.54	2.02
					2.44	2.55
						2.52

Standardized Euclidean

3.31	2.69	3.07	1.58	3.32	3.29	1.59
	2.53	2.67	3.24	2.68	2.47	3.29
		2.70	2.70	2.95	3.27	3.02
			3.02	2.67	2.50	2.84
				3.34	3.22	2.09
					2.74	3.09
						3.05

Minkowski p=0,5

2923.39	1449.41	1931.58	194.60	3366.42	3701.25	146.93
	1558.87	1549.31	2266.13	2232.32	1585.78	2847.43
		1582.30	1129.40	3054.05	3531.18	2095.64
			1855.12	2000.59	1860.27	1605.69
				2824.09	3476.19	532.27
					2513.10	2736.72
						2885.47