

A Quality Model for Spreadsheets

Jácome Cunha, João Paulo Fernandes, Christophe Peixoto, João Saraiva
HASLab, INESC TEC & Universidade do Minho, Portugal
{jacome,jpaulo,christophe,jas}@di.uminho.pt

Abstract—In this paper we present a quality model for spreadsheets, based on the ISO/IEC 9126 standard that defines a generic quality model for software. To each of the software characteristics defined in the ISO/IEC 9126, we associate an equivalent spreadsheet characteristic. Then, we propose a set of spreadsheet specific metrics to assess the quality of a spreadsheet in each of the defined characteristics.

In order to obtain the normal distribution of expected values for a spreadsheet in each of the metrics that we propose, we have executed them against all spreadsheets in the large and widely used EUSES spreadsheet corpus. Then, we quantify each characteristic of our quality model after computing the values of our metrics, and we define quality scores for the different ranges of values.

Finally, to automate the attribution of a quality score to a given spreadsheet, according to our quality model, we have integrated the computation of the metrics it includes in both a batch and a web-based tool.

I. INTRODUCTION

Spreadsheets are widely used both by professional and non-professional programmers. Non-professional programmers see them as simple, flexible and easy-to-use calculators, whereas professional ones exploit them in software intensive organizations as advanced intermediate representations to perform data migration between software systems and to perform operations to enrich or simplify data, for example.

The simplicity of spreadsheets and their attractive visual representation has ensured spreadsheet systems a great popularity, being them nowadays one of the most used software systems in the world. Unfortunately, spreadsheets are known for being error prone, as reported by numerous studies which show that up to 90% of real-world spreadsheets contain errors [1], [2].

While the amount of errors found in spreadsheets suggests that their quality is inherently low, the fact is that, surprisingly, with the exception of the works [3] and [4] by Correia and Ferreira at the Software Improvement Group (SIG)¹, little work has been done on trying to assess the quality of general purpose spreadsheets. As a result, there are no methods, techniques nor tools to measure and, consequently, to improve the quality of a spreadsheet.

In this paper we propose a quality model for spreadsheets. This model is defined by introducing a set of domain specific

This work is funded by the ERDF through the Programme COMPETE and by the Portuguese Government through FCT - Foundation for Science and Technology, project ref. PTDC/EIA-CCO/108613/2008. The three first authors were also supported by FCT grants SFRH/BPD/73358/2010, SFRH/BPD/46987/2008 and B13-2011PTDC/EIA-CCO/108613/2008, respectively.

¹Software Improvement Group (www.sig.eu).

metrics for spreadsheets which are used to measure concrete spreadsheet characteristics. Such characteristics are related to the general notions of software quality as defined on the ISO/IEC 9126 standard [5]. To calibrate our quality model we use a large repository of real-world spreadsheet instances, namely the EUSES Spreadsheet corpus [6].

In order to automate the assessment of the quality of a spreadsheet, we have implemented our quality model in a (batch) software tool which, together with the EUSES corpus, allows us to automatically study the behaviour of the spreadsheet metrics, to give quality scores to spreadsheets, and to evolve our quality model.

This paper is organized as follows: in section II we briefly present the ISO/IEC 9126 software quality standard and how we instantiate it to the spreadsheet realm. In section III we build on the ISO/IEC 9126 to propose a spreadsheet quality model. The calibration of the proposed model is discussed in section IV, and the model itself is evaluated in section V. In section VI we compare our work with works whose goals are related to ours and in section VII we conclude the paper.

II. A QUALITY MODEL FOR SPREADSHEETS

We define a quality model for spreadsheets based on the widely accepted ISO/IEC 9126 international standard for software product quality [5]. This standard provides a terminology regarding the concept of software product quality that distinguishes six main characteristics. These characteristics are then sub-divided in sub-characteristics as shown in Figure 1.

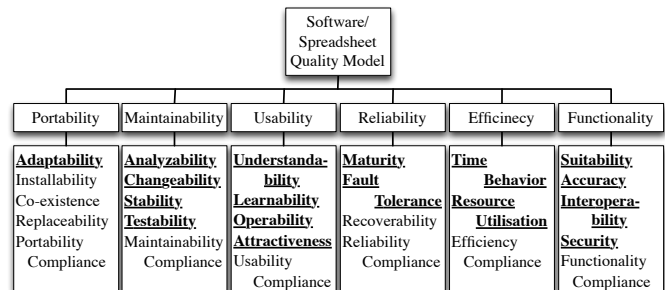


Fig. 1: The ISO/IEC 9126 standard for software quality.

In order to define our quality model for spreadsheets we follow a three phase approach:

- 1) Firstly, we instantiate the generic ISO/IEC 9126 characteristics to the context of spreadsheets. That is to say that, whenever possible, we define domain specific spreadsheet characteristics that directly relate to the

more general ISO/IEC 9126 characteristics. Because the ISO/IEC 9126 was created with the goal of assessing the quality of all software products, there are some characteristics of the model that do not directly apply to spreadsheets. This is the case, for example, of the recoverability characteristic that is very important in most software products, but not in regular spreadsheet systems. In Figure 1 the sub-characteristics shown in bold are the ones we directly map into spreadsheet equivalents.

Moreover, we define a set of spreadsheet metrics that allow us to assess the quality of those characteristics.

- 2) Secondly, in order to calibrate our quality model we consider a large repository of real-world spreadsheets. For this, we use the EUSES spreadsheet corpus, a shared resource for supporting experimentation with spreadsheets [6]. This corpus contains a large sample of spreadsheets suitable for automated processing. The calibration of the quality model is then performed by computing the metrics for all spreadsheets in the repository. Then, and based on this global analysis of the complete repository, we define for each metric used in each characteristic, six intervals of possible metric values. These intervals are used to define a five star ranking as it is currently widely used on software product markets. Moreover, by combining different characteristic scores/stars we are able to give quality scores/stars to spreadsheets.
- 3) Thirdly, to evaluate and further calibrate our quality model we manually assess the quality of several EUSES spreadsheets, and we compare these results to the automatically generated ones. This manual evaluation allowed us to evolve our quality model.

The calibration of our model requires that we process a large number of spreadsheets. Thus, we have developed a tool that computes several metrics given a spreadsheet. The three phases of our approach are described in the next three sections.

III. THE ISO/IEC 9126 FOR SPREADSHEETS

In this section we discuss each of the ISO/IEC 9126 characteristics in the context of spreadsheets. We briefly describe those characteristics in the context of spreadsheets, and we discuss spreadsheet metrics we may use to assess the quality of such characteristics.

Functionality: Functionality is the capacity of spreadsheets to satisfy the needs of their end users, either implied or stated. It is divided by the following four sub-characteristics:

- 1) *Suitability:* is the quality of a spreadsheet having the right properties for a specific purpose. To assess the quality of this characteristic we consider that if a spreadsheet has formulas with references to blank cells then it does not fit the desired property. Similarly, spreadsheet incongruences may also affect suitability. By a spreadsheet incongruence we mean, for example, a cell whose

type deviates from a pattern: when all but one cells in a same column are defined by a consistent formula, which is defined by a constant value.

- 2) *Accuracy:* is the faithful measurement or representation of the correctness of a spreadsheet. A great number of both incongruences and blank cells referenced in formulas would drop the accuracy of a spreadsheet. In the same line, if a spreadsheet has many output cells containing errors, then its accuracy will be greatly affected.
- 3) *Interoperability:* is the ability of two or more spreadsheet worksheets or components to exchange information and to use the information that has been exchanged. We can say that a spreadsheet has a good interoperability if most of its formulas are correct since that means that the components are changing information without any problem. If the spreadsheet has many references, cells with references and data been exchanged between its worksheets then it has a good interoperability.
- 4) *Security:* is the ability of a spreadsheet to be protected. Since spreadsheets are mostly used by end-users with little knowledge of programming, a spreadsheet is safe if some of its cells like the formulas or Data are protected, preventing the user to change them by mistake. Besides that, the spreadsheet security can be increase with a Password locking the Workbook or the Worksheets, preventing anyone unauthorized to access or change it.

Reliability: Reliability is the capacity of a spreadsheet to maintain its level of performance under stated conditions for a stated period of time. It is divided in the following two sub-characteristics:

- 1) *Maturity:* The state of quality of a spreadsheet being fully developed. If the spreadsheet contains non used worksheets, then we consider that it may not be completely developed. The existence of empty labels in rows or columns, blank cells, or tables with blank cells may also indicate that the spreadsheet is under development.
- Fault Tolerance:* It is the property of a spreadsheet to continue operating properly in the event of one or more faults within some of its components.

If a cell is referenced by many other cells (directly or indirectly), lesser is the fault tolerance since modifying that cell can diffuse mistakes on all the cells that reference it. Furthermore, if a spreadsheet contains many complex formulas, is less fault tolerant because those formulas reference many cells, so changing just one of those can lead to an error on the formula result

Usability: Usability is the capacity of the spreadsheet to be understood, learning how it works, be used and intuitive to the user. It is divided by the following four sub-characteristics:

- 1) *Understandability:* It is capacity of being understood. There is many ways to make a spreadsheet more understandable, being the more important one separate the Input, the Computation and Output parts of the spreadsheet. Also giving different background colors for

the different types of Data cells can turn the spreadsheet easier to understand. Finally, having a big number of cells makes the spreadsheet harder to understand

- 2) *Learnability*: It is the capacity to enable end users to use the spreadsheet. The more cells and complex formulas a spreadsheet have, more difficult is to learn how to use it. The same if the spreadsheet have many references and Data been exchanged between its sheets. On the other hand to improve its learnability we can give different background colors for different types of Data, and separate the Input, Computation and Output parts.
- 3) *Operability*: It is the capacity of the spreadsheet being operated. If we separate the Input from both the Computation and Output then the spreadsheet is easier to operate. If the spreadsheet has data validation drop down lists, it would also be easier to operate.
- 4) *Attractiveness*: It is the capacity to be attractive/appealing to the end user. As on the others sub-characteristics from the Usability, it is important to have the Input, Computation and Output parts of the spreadsheets separate from each other in order to be more appealing for the user. Besides that by giving different background colors for different types of Data cells and having data validation drop down lists, also improve the attractiveness.

Efficiency: Efficiency is the amounts of resources used are compatible with the performance level of the spreadsheet. It is divided by the following two sub-categories:

- 1) *Time Behavior*: Both search formulas and vlookup functions are operations that may require intensive computation, so the of such operations can influence the time behaviour of a spreadsheet. Moreover, if the spreadsheet has many complex formulas, it will also increase the computing time.
- 2) *Resources Utilization*: Since the vlookup function use many resources, the number of vlookup functions influences the resources needed. Furthermore, if the spreadsheet has a high number of non-blank cells or complex formulas, is going to use more resources.

Maintainability: Maintainability is the capacity of a spreadsheet to be modified, either to expand functionalities or to correct errors. It is divided by the following four sub-characteristics:

- 1) *Analyzability*: It is the capacity to be analyze, to conclude the effort needed for diagnosis deficiencies. If a spreadsheet has a big number of cells, References or formulas, is harder to analyze and find the possible deficiencies. A good way to turn the spreadsheet easier to analyze is to have the Data well organized (separate Input, Computation and Output)
- 2) *Changeability*: How well can a spreadsheet be changed, conclude the effort needed for modifications, fault removal. A spreadsheet with well-organized Data is easier to change and to remove faulty parts. On the other side

if the spreadsheet has a great number of cells or cells Referenced is harder to change.

- 3) *Stability*: It is the capacity to be stable. A spreadsheet lose stability if it has a big number of cells been referenced by other cells, since changing that first cell can spread errors all over the spreadsheet. It also lose stability if it has many Complex formulas, since changing just one cell from the many referenced by the complex formula can change the formula result in to a wrong one.
- 4) *Testability*: It define how well can the SpreadSheet be tested. The relevance on testing is to confirm that the result we have is the expected one, on spreadsheets we need only to check the Output cells, that can be formula cells or Data cells, but those last ones are usually labels so we just need to test the formula cells. So bigger is the number of formula cells, harder it is to test the spreadsheet.

Portability: Portability is the capacity to be transferred from one environment to another. It is divided by the following sub-characteristic:

- 1) *Adaptability*: it defines how well can the spreadsheet be adapted to environmental change. Since macros are not compatible with all the environment or spreadsheet applications, the higher number of Macros on a spreadsheet less adaptable that spreadsheet is.

IV. USING METRICS TO CALIBRATE THE MODEL

In the previous section we have defined the ISO/IEC 9126 characteristics in the context of spreadsheets. We have also discussed spreadsheet metrics which can be used to assess quality of those characteristics. In this section we study how to calibrate our quality model based on spreadsheet metrics.

To calibrate our model we need to consider a large repository of spreadsheets. The idea is to compute the metrics in a large set of real-word spreadsheet in order to define metric values that allow us to assign a five star ranking to a spreadsheet. Thus, we consider the large EUSES spreadsheet corpus that consists of 5607 spreadsheet files, classified according six categories, namely, financial, inventory, homework, grades, database and modeling, as shown in Figure 2.

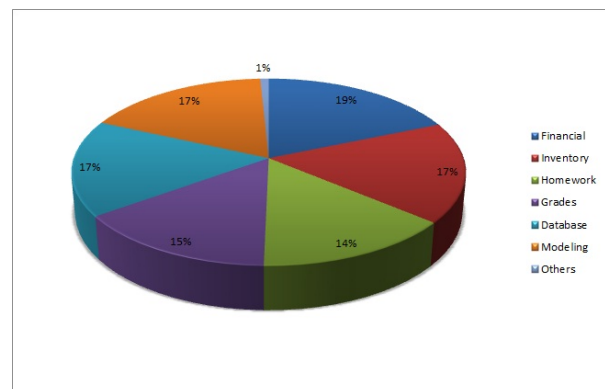


Fig. 2: Percentage of spreadsheets in the EUSES by category.

The calibration of the quality model is then performed by computing the metrics for all spreadsheets in the EUSES repository. Then, and based on this global analysis of the complete repository, we define for each metric used in each characteristic, six intervals of possible metric values. There are several techniques used to aggregate metrics to ratings [3]. In our quality model we consider the central tendency using the median as suggested in [7]. Figure 3 displays the five graphics that represent the aggregation based on the central tendency of the results of computing the total cells metric, for the complete EUSES spreadsheet repository, in a five level rating.

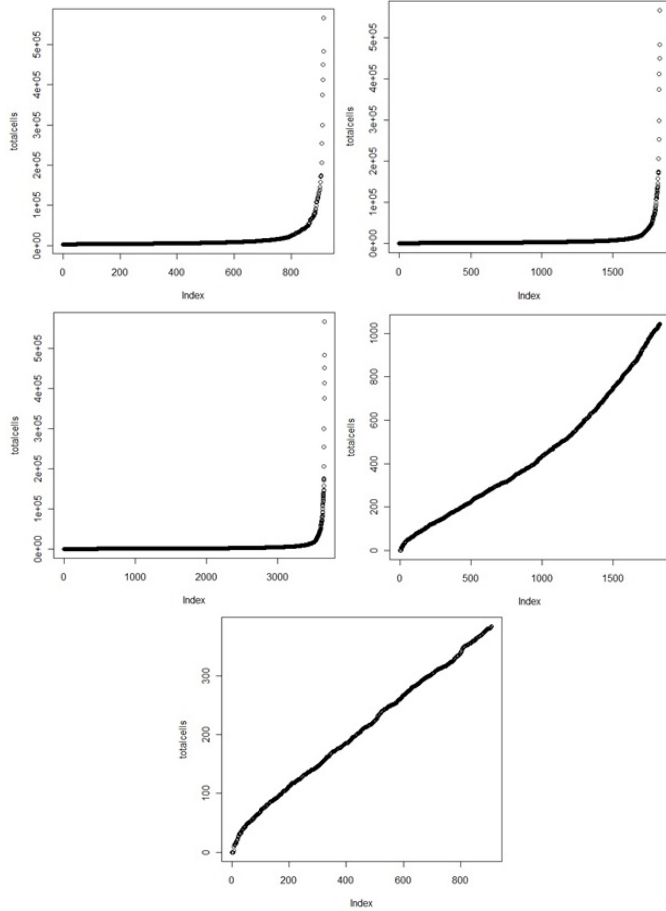


Fig. 3: Graphics of the analysis of the total cells metric.

These intervals are used to define a five star ranking, where 0 star represent the lowest quality ranking and five stars the higher quality ranking, as it is currently widely used on software product markets. Table I presents for each characteristic considered in our quality model, the metric values intervals defining the five star rating.

V. MODEL EVALUATION

In this section we present the evaluation we performed to assess the sharpness of the results produced by our model.

We have implemented our quality model in two different tools: a batch and a web-based spreadsheet quality assessment tools. The batch tool can be used to analyze regular

spreadsheets and the web-based version to assess the quality of spreadsheet within the GoogleDocs environment.

For each EUSES category, we have chosen in a random manner five spreadsheets whose quality was determined automatically by our (batch) tool. In Figure 4 we present the results we obtained when assessing the quality of five spreadsheets in the *database* category: for each of the characteristics that we consider a star-based mark is assigned.

| Characteristics - Files | os_content_inventory | Budget | CollectionsReports | Basicdata | centermaterials |
|-------------------------|----------------------|--------|--------------------|-----------|-----------------|
| Functionality | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Suitability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Accuracy | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Interoperability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Reliability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Maturity | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Fault Tolerance | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Usability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Understandability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Learnability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Operability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Attractiveness | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Efficiency | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Time behavior | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Resource utilisation | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Maintainability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Analysability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Changeability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Stability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |
| Testability | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ | ★★★★★ |

Fig. 4: Database stars.

Using our batch tool we were able to assess the quality of the complete EUSES. In order to assess the appropriateness of our tool (and of our model) we have manually inspected 30 EUSES spreadsheets and we have compared the grades we would manually attribute to a spreadsheet against the grades automatically calculated by our tool. We report in the next paragraphs the results of this exhaustive work; our analysis follows the categories defined within EUSES.

a) *Database Category*: The sub-characteristics *suitability* and *accuracy* of *functionality* were not always correctly computed. In our analyzes it seems that one of the analyzed spreadsheets was being used as a text document instead of a spreadsheet. This has lead to imprecise automatic results. Nevertheless, we have not found any other similar occurrence. The *maturity* of *reliability* were also inappropriately calculated for two spreadsheets. This was due to the fact the algorithm that calculates the empty-cells metric considers as empty cells the rows/columns with no data separating different parts of the spreadsheet. In fact, this is a problem reported in other contexts [8]. In this case, the algorithm should be improved to discard the empty cells that are used to improve the layout of the spreadsheet. Finally, the *attractiveness* of *usability* was also calculated incorrectly for one spreadsheet. The reason for this was that the current metrics do not give information about the organization of the data, that is, its *layout* (e.g., different colors in the labels). We believe, however, that to improve the usability characteristic star rating requires extremely complex metrics, and this is a direction that we have not yet explored.

b) *Modeling Category*: The *maturity* sub-characteristic of *reliability* was miss calculated for one spreadsheet. The reason remains the the same as before: the empty-cells metric still can be improved. The *understandability* of *usability*, for one spreadsheet was imprecisely rated, again due to the nature

| Metric | 5 stars | 4 stars | 3 stars | 2 stars | 1 stars | 0 stars |
|---|---------|-----------|------------|--------------|---------------|---------|
| Functionality | | | | | | |
| Suitability | | | | | | |
| number of incongruences | [0,2] |]2,9] |]9,18] |]18,62] |]62,141] | >141 |
| number of blank cells referenced in formulas | [0,4] |]4,12] |]12,55] |]55,266] |]266,710] | >710 |
| Accuracy | | | | | | |
| number of incongruences | [0,2] |]2,9] |]9,18] |]18,62] |]62,141] | >141 |
| number of blank cells referenced in formulas | [0,4] |]4,12] |]12,55] |]55,266] |]266,710] | >710 |
| number of output cells with errors/bad content | | | | | | |
| Interoperability | | | | | | |
| number of cells that have references | [0,2] |]2,12] |]12,63] |]63,187] |]187,416.5] | >416.5 |
| number of references | [0,33] |]33,80] |]80,304] |]304,1141] |]1141,2974] | >2975 |
| data been exchanged between sheets; number of correct formulas | | | | | | |
| Security | | | | | | |
| number of protected cells for writing; password to lock data; hidden information; data validity constraints | | | | | | |
| Reliability | | | | | | |
| Maturity | | | | | | |
| difference between the nr. of existing sheets and the nr. of sheets used | 0 | 1 | 2 |]2,9] |]9,15] | >15 |
| number of blank cells | [0,8] |]8,60] |]60,320] |]320,1271.5] |]1271.5,3341] | >3341 |
| number of labeled rows/columns that are empty | | | | | | |
| Fault Tolerance | | | | | | |
| number of referenced cells | [0,33] |]33,80] |]80,304] |]304,1141] |]1141,2974] | >2975 |
| number of complex formulas | [0,1] |]1,4] |]4,18] |]18,66] |]66,162] | >162 |
| Usability | | | | | | |
| Understandability | | | | | | |
| number of cells | [0,208] |]208,385] |]385,1045] |]1045,2858] |]2858,5794] | >5794 |
| different colors for different types of data; separate input, computation and output | | | | | | |
| Learnability | | | | | | |
| number of complex formulas | [0,1] |]1,4] |]4,18] |]18,66] |]66,162] | >162 |
| number of cells | [0,208] |]208,385] |]385,1045] |]1045,2858] |]2858,5794] | >5794 |
| number of references | [0,33] |]33,80] |]80,304] |]304,1141] |]1141,2974] | >2974 |
| different colors for different types of data; separate input, computation and output; amount of data being exchanged between sheets | | | | | | |
| Operability | | | | | | |
| number of referenced cells | [0,2] |]2,12] |]12,63] |]63,187] |]187,416.5] | >416.5 |
| create/have data validation drop down lists; separate input, computation and output | | | | | | |
| Attractiveness | | | | | | |
| number of non-blank columns | [0,3] |]3,5] |]5,8] |]8,12] |]12,18] | >18 |
| number of cells | [0,208] |]208,385] |]385,1045] |]1045,2858] |]2858,5794] | >5794 |
| create/have data validation drop down lists; different colors for different types of data; separate input, computation and output | | | | | | |
| Efficiency | | | | | | |
| Time Behavior | | | | | | |
| number of complex formula | [0,1] |]1,4] |]4,18] |]18,66] |]66,162] | >162 |
| Time Behavior | | | | | | |
| number of non-blank cells | [0,81] |]81,162] |]162,412] |]412,1125] |]1125,2196] | >2196 |
| number of formulas | [0,5] |]5,16] |]16,65] |]65,235] |]235,540] | >540 |
| Maintainability | | | | | | |
| Analyzability | | | | | | |
| number of cells | [0,208] |]208,385] |]385,1045] |]1045,2858] |]2858,5794] | >5794 |
| number of formulas | [0,5] |]5,16] |]16,65] |]65,235] |]235,540] | >540 |
| number of references | [0,33] |]33,80] |]80,304] |]304,1141] |]1141,2974] | >2974 |
| data organization | | | | | | |
| Changeability | | | | | | |
| number of cells | [0,208] |]208,385] |]385,1045] |]1045,2858] |]2858,5794] | >5794 |
| number of referenced cells | [0,2] |]2,12] |]12,63] |]63,187] |]187,416.5] | >416.5 |
| data organization | | | | | | |
| Stability | | | | | | |
| number of complex formulas | [0,1] |]1,4] |]4,18] |]18,66] |]66,162] | >162 |
| number of referenced cells | [0,2] |]2,12] |]12,63] |]63,187] |]187,416.5] | >416.5 |
| Testability | | | | | | |
| number of formulas | [0,5] |]5,16] |]16,65] |]65,235] |]235,540] | >540 |
| Portability | | | | | | |
| Adaptability | | | | | | |
| quantity of macros code | | | | | | |

TABLE I: Calibration of the quality model based on the metrics calculated on the EUSES spreadsheets.

of the empty-cells metric. Finally, the *understandability* and *attractiveness* of *usability* were imprecisely calculated because we have not considered layout metrics.

c) *Homework Category*: The *maturity* sub-characteristic of *reliability* was miss calculated for one spreadsheet because of the empty-cells metric algorithm. The *understandability* and *attractiveness* of *usability*, for one spreadsheet, were also inaccurately calculated because no layout information is calculated.

d) *Inventory Category*: The *understandability* and *attractiveness* sub-characteristics of *usability*, for one spreadsheet were imprecisely calculated again because no layout information is exploited.

For the remaining (sub)characteristics we have observed that the computed stars are in fact precise and accurate.

For each spreadsheet a total amount of 20 ratings (number of stars) is calculated (only for sub-characteristics and characteristics). Thus, for the 30 spreadsheets analyzed, our tool computed 600 ratings. From our manual inspection, we have only observed 24 imprecise star-ratings. This means that our tool rated accurately 96% of the characteristics and sub-characteristics. From this exhaustive analyze, we believe it is appropriate to conclude that in general our tool is producing good results. Nevertheless, we spotted some limitations:

- In spreadsheets empty cells are very often used with layout purposes (very much as spaces are used in regular programming languages). The software used to compute the empty cells metric does not handle layout information, so such “layout” empty cells are considered! We may use Erwig’s spacial logic algorithms to have a more precise metric for empty cells [9], [10].
- Our software does not consider the use of colors, tables, and other layout information. Thus, more non-trivial metrics must be added to the algorithm so more precise quality ratings can be calculated.

This is the very first iteration in the goal of defining a quality model for spreadsheets. This model needs now to be evolved, both by adding new metrics to the model and by defining new intervals for the star ranking. In fact, from these evaluation, we can conclude that our spreadsheet quality model can be improved: the sub-characteristics *understandability* and *attractiveness* should be enhanced with a new metric to calculate the quality of the spreadsheet layout. This would improve the quality of the rating computed by our model to the *usability* characteristic, which is the one that was more miss calculated.

VI. RELATED WORK

In [4] the authors take a first step towards automated assessment of spreadsheet maintainability. As in the work here presented, they apply the selected metrics to the EUSES spreadsheet corpus in order to study their behavior. Their work, however, is restricted to achieve a maintainability model whilst we defined a first complete spreadsheet quality model.

In [11] the authors sketch a new maintainability model that alleviates some problems reported by other techniques. Since

this work was done in an industrial environment, they discuss their experiences with using such a model for IT management consultancy activities. Although their model features only maintainability of software, their technique to calibrate the model is similar to the one we used.

Metrics for spreadsheets have already been defined [12], [13]. Nevertheless, to the best of our knowledge, our work is the very first attempt to define a complete spreadsheet quality model. Although this is a first proposal, we have already defined it based on the general software quality model, calibrated it based the most used (in science) set of spreadsheets and manually validated it (been suggested some improvements).

VII. CONCLUSIONS

This paper presented a quality model for spreadsheets based on the ISO/IEC 9126 standard. Spreadsheet characteristics have been considered to model the ISO/IEC 9126 standard, and spreadsheet metrics have been defined to assess the quality of such characteristics. We have calibrated and validated our quality model using the large EUSES spreadsheet repository. Finally, we have implemented our quality model in a software tool and our first experimental results show the tool is able to give good results automatically assessing the quality of a spreadsheet.

REFERENCES

- [1] R. R. Panko, “Spreadsheet errors: What we know. what we think we can do,” *Symp. of the European Spreadsheet Risks Interest Group (EuSpRIG)*, 2000.
- [2] R. Panko, “Facing the problem of spreadsheet errors,” *Decision Line*, 37(5), 2006.
- [3] J. P. Correia and M. A. Ferreira, “Requirements for automated assessment of spreadsheet maintainability,” *CoRR*, vol. abs/1111.6902, 2011.
- [4] —, “Measuring maintainability of spreadsheets in the wild,” in *IEEE 27th Int. Conf. on Software Maintenance (ICSM)*, 2011, pp. 516–519.
- [5] ISO(2001), “ISO/IEC 9126-1: Software engineering - product quality - part 1: Quality model,” Geneva, Switzerland.
- [6] M. F. II and G. Rothermel, “The EUSES Spreadsheet Corpus: A shared resource for supporting experimentation with spreadsheet dependability mechanisms,” in *Proceedings of the 1st Workshop on End-User Software Engineering*, 2005, pp. 47–51.
- [7] D. Spinellis, “A tale of four kernels,” in *Proceedings of the 30th international conference on Software engineering*, ser. ICSE ’08. New York, NY, USA: ACM, 2008, pp. 381–390.
- [8] J. Cunha, J. P. Fernandes, H. Ribeiro, and J. Saraiva, “Towards a catalog of spreadsheet smells,” in *Proceedings of The 12th Int. Conference on Computational Science and Its Applications*, 2012, to appear.
- [9] R. Abraham and M. Erwig, “Header and unit inference for spreadsheets through spatial analyses,” *Visual Languages and Human Centric Computing, 2004 IEEE Symposium on*, pp. 165–172, Sept. 2004.
- [10] —, “UCheck: A spreadsheet type checker for end users.” *J. Vis. Lang. Comput.*, vol. 18, no. 1, pp. 71–95, 2007.
- [11] I. Heitlager, T. Kuipers, and J. Visser, “A practical model for measuring maintainability,” in *Proceedings of the 6th International Conference on Quality of Information and Communications Technology*, ser. QUATIC ’07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 30–39.
- [12] A. Bregar, “Complexity metrics for spreadsheet models,” *CoRR*, vol. abs/0802.3895, 2008.
- [13] K. Hodnigg and R. T. Mittermeir, “Metrics-based spreadsheet visualization: Support for focused maintenance,” *CoRR*, vol. abs/0809.3009, 2008.