Metrology and Quantitative Analysis in ISO 15939

Alain Abran¹, Jean-Marc Desharnais¹, and Juan Jose Cuadrado-Gallego²
¹École de Technologies supérieure – Université du Québec, Montréal, Canada
²Universidad de Alcala de Henares, Madrid Spain

Abstract – Measurement based on international standards for measurement (i.e. metrology) is not the same as the judgment-based quantification of implicit relationships across a mix of entities and attributes without due consideration of admissible mathematical operations on numbers of different scale types. This paper analyzes the Measurement Information Model in ISO 15939 and clarifies what in it refers to the classical metrology field, and what refers to the quantitative analysis of relationships.

Keywords: Measurement, Metrology, ISO 15939, ISO VIM

1 Introduction: Numbers are not all created equal

Software practitioners and researchers alike often forget that numbers are not all created equal. For instance, a number derived from the result of a measurement process which meets the metrology requirements is a quantity expressed with a measurement unit. By contrast, a number derived from a mix of mathematical operations without consideration of measurement units and scale types will still be a number, but it could be a meaningless one (some of the Halstead’s software metrics, for example [1]). Practitioners may feel good about such a potpourri of numbers from models which appear to take into account a large number of factors (i.e. as in many estimation and quality models) – see, for example, Use Case Points [2]. However, feeling good does not add validity to mathematical operations that are inadmissible in measurement.

In practice, various types of quantitative models produce numbers in outputs (i.e. the outcomes of the models), but they do not have the same qualities as the numbers that meet the requirements of metrology:

• An estimation model will provide a number as an estimate. However, with every such estimated number is associated a (potentially large) range of variations, depending on the number of input parameters and their corresponding uncertainties, as well as on the uncertainties about the relationships across all such parameters. The estimated numbers are not meaningful without a knowledge (and understanding) of the corresponding uncertainties.

• A quality model will provide a number which typically depends on: a specific selection among a (potentially large) number of alternatives; the assignment of a percentage to each contributing alternative, based on the opinion of one person (or a group of persons); and comparison of the contributing alternatives, with their distinct threshold values that are often defined by opinion as well.

In many instances, in these analysis models,

• some, if not all, of the numbers used as inputs to them are obtained by opinion, rather than from precise measurements (i.e. with measurement instruments or from the application of detailed measurement procedures);

• the input numbers are combined without explicitly describing the admissible mathematical operations and the treatment of the corresponding measurement units; and

• while the outputs of such models are indeed numbers, they do not have metrological properties, and should be handled very cautiously.

Analysis models like these are quantitative models, but they are not measurement models in the metrological sense. Such differences between quantitative analysis and measurement are not generally discussed in the software engineering literature. In this paper, we discuss these differences, using in particular the ISO 15939 Measurement Information Model which contains a metrology-related perspective as well as an analysis perspective. These concepts are illustrated with a productivity model as a Measurement Information Model.

2 ISO 15939 Measurement Information Model

The Measurement Information Model from ISO 19539 [3] (Figure 1) sets out the various steps necessary for the design of an information product when a measurable concept has to be designed and used in practice. In the illustration of this model in the figure, ovals represent activities and rectangles represent the
input and output of an activity. Figure 1, when read from the bottom up, shows the following:

1. A specific measurement method has to be designed to obtain a base measure for a specific attribute;
2. The values of two or more base measures can be used in a computational formula (by means of a measurement function) to construct a specific derived measure;
3. Derived measures are used in the context of an analysis model of relationships to construct an indicator;
4. The indicator is used for interpretation purposes to build the information product to meet the information needs. This means that the indicator’s value is interpreted within the prescribed context as describing, in the language of the measurement user, an information product for his information needs [3].

It is to be noted that, in 2 and 3 above, both the derived measures and the indicator inherit the properties of the mathematical operations upon which they are built:
• These numbers are meaningful (i.e. have valid properties) when they are derived from admissible mathematical operations.
• These numbers are meaningless when they are derived from inadmissible mathematical operations, or when the measurement units and measurement scale types are not considered correctly within the mathematical operations.

3 Scope of the ISO 15939 Measurement Information Model

To better understand the Measurement Information Model in ISO 15939, it is useful to identify what in the model is related to metrology concepts and what is not.

A- Metrology-related: measurement of the attribute of an entity

The bottom portion of the ISO 15939 Measurement Information Model can be mapped to the metrology concepts [4-7] in two steps – see Figure 2:
1. Data collection: when a measurement method is used to measure an attribute, the output is the base measure of each specific entity being measured – this corresponds to the data collection step of mapping;
2. Data preparation: when a number of the base measures of the data collected are combined through a measurement function (using an agreed-upon mathematical formula and related labels), the combined units are considered as derived measures – this corresponds to the data preparation step of mapping, prior to the analysis phase.

B- Non Metrology-related: quantification of relationships across attributes and entities.

The top portion of the ISO 15939 Measurement Information Model deals with the analysis (through quantification) of relationships across entities and attributes. This analysis part of the ISO 15939 Measurement Information Model includes two activities (i.e. the ovals in the figure):
1. Analysis model: modeling of the relationships across entities and attributes to derive an indicator of the value of such relationships.
2. Interpretation: interpretation of the indicator to produce the information product that is typically used in an evaluation or decision making process.

The metrology-related bottom part of the Measurement Information Model is supported by the set of metrology concepts described in [8-9]. The upper part of the Measurement Information Model is outside the scope of metrology [3], since it deals with the use of the measurement results from the lower part of the model. This analysis is not extensively described in ISO 19539, except through a few complex definitions tailored to that specific standard. These two perspectives are discussed in greater detail in the following two sections.

4 The Metrology Perspective in the ISO 15939.

This section describes the metrology-related part of the ISO 15939 Measurement Information Model – i.e. the bottom part of Figure 2.

4.1 Data collection: base measures

Every base measure must correspond to a single distinct software attribute (i.e. a property of an object or concept). So, identifying the attribute of the entity to be measured and quantifying it through its measurement method constitutes the data collection step. Of course, the attribute must be well defined; if it is not, designing an adequate measurement method would be pretty challenging.

4.2 Data preparation: derived measures

Depending on the information needs, some of the base measures already collected for an entity can be assembled in accordance with a measurement function (e.g. a computational formula) defined for each derived measure – see Figure 2 in the Data Preparation step. A derived measure is therefore the product of a set of measurement units properly combined (through a measurement function).

If a derived measure is designed in bottom-up fashion, the name assigned to this combination of units should correspond to the concept representing the particular combination of measurable attributes. The
accuracy of a derived measure (along with the corresponding measurement errors) is directly related to the accuracy of each of its base measures and to how these base measures are mathematically combined. Stated differently, the qualities of the corresponding measuring instrument(s) of the base measures impact the quality of the derived measures. For example, the accuracy of a velocity measurement will directly depend on the accuracy of its two base measures: distance and time.

When their corresponding base measures are not sufficiently well defined, standardized, and instrumented to ensure the accuracy, repeatability, and repetitiveness of measurement results, then, when the same entity (software) is measured by different measurers, the results can be significantly different. It must be noted that a derived measure is descriptive. It does not explain a relationship, nor does it say anything about the strength of such a relationship.

Example of a derived measure: velocity. The combination distance traveled over a period of time (e.g. km/hour) is associated with the concept of velocity. Such a derived measure (i.e. velocity) is typically measured by a measuring instrument which:

- captures both base measures simultaneously (that is, distance and time, measured in meters and seconds on a car’s speedometer, for instance),
- has an integrating feature which divides the base measures to produce a ratio (time/distance) to represent the velocity concept, and
- has a display feature which shows up the measurement results using a standardized display convention: for example, converting meters per second into the universally adopted standard for cars, which is kilometers per hour.

5 The Quantification of Relationships in ISO 15939

This section looks at the Analysis of Relationships in the ISO 15939 Measurement Information Model – the top part of Figure 2.

5.1 Quantitative Elements of the ISO 15939 Analysis Model

The top part of the ISO 15939 Measurement Information Model deals with the use of measurement results in various evaluation or decision making models. This use of measurement results (which have been obtained from a metrology-based approach) is represented very succinctly in ISO 15939 using only:

- two activities (the ovals in Figures 1 and 2): (Analysis) Mode and Interpretation;
- one number (the rectangles in Figure 1 and 2): Indicator.

In practice, this use of measurement results typically involves:

- analysis of the relationships across different measurement results with respect to various conditions within a context, and
- assessment against reference contexts for evaluation and/or decision making.

The intricacies and subtleties of the above tasks are not graphically represented in the Measurement Information Model in Figure 1, but are to be found in the textual descriptions of the following three expressions in ISO 15939: Indicator, (Analysis) Model, and Decision criteria, where some of the terms not represented in the model are underlined.
A number of the concepts presented in these descriptions do not appear in the Measurement Information Model of ISO 15939, such as: Decision criteria, Assumptions, Expected relationships, Estimates or evaluation, Numerical thresholds or targets, Statistical confidence limits, etc.

5.2 Refined representation of the Analysis Model

We pointed out in the previous subsection that a number of the concepts mentioned in the three descriptions in the side box are not directly modeled in Figure 1. To facilitate an understanding of the relationships across the many concepts embedded within the ISO 15939 Measurement Information Model, the set of key concepts has been extracted from these three descriptions and modeled in Figure 3.

Figure 3: Refined Analysis Model of the ISO 15939 Measurement Information Model

The refined representation of the Analysis Model represented in Figure 3 includes two additional major blocks:

- A standard reference model (Figure 3, bottom left), which can include, for instance, an accepted model of the relationships across distinct types of objects of interest. When such a reference model exists, it can be: an industry model, an ISO model, or a generally accepted statistical technique (and related mathematical model). This standard reference model would include: the set of formal (or informal and assumed) individual relationships, along with the base or derived measures to be considered as evaluation or decision criteria, as well as the algorithm (mathematical or implied) that combines them in a criterion.

- An organizational reference context (Figure 3, upper left), ideally aligned with the standard reference, with a set of selection criteria and values specific to the organization. This organizational reference would contain the reference values necessary for interpretation, including: A) a set of reference values specified for this context; and B) evaluation or decision criteria with either target values or specific evaluation scales.

The ISO 15939 Definitions for the Use of Measurements Results

**Indicator:** An indicator is a measure providing an estimate or evaluation of specified attributes derived from a model with respect to defined information needs. Indicators are the basis for analysis and decision making. These are what should be presented to measurement users.

**Analysis Model:** An algorithm or calculation combining one or more base and/or derived measures with associated decision criteria. It is based on an understanding of, or assumptions about, the expected relationship between the component measures and/or their behavior over time. Models produce estimates or evaluations relevant to defined information needs. The scale and measurement method affect the choice of analysis techniques or models used to produce indicators.

**Decision criteria:** Decision criteria are numerical thresholds or targets used to determine the need for action or further investigation, or to describe the level of confidence in a given result. Decision criteria help to interpret the results of measurement. Decision criteria may be calculated or based on a conceptual understanding of expected behavior. Decision criteria may be derived from historical data, plans, and heuristics, or computed as statistical control limits or statistical confidence limits.

NOTE: Some of the terms not represented in the model in Fig. 1 are underlined above in the descriptions.

5.3 The (Implicit) Link between Measurement and Quantitative Analysis of Relationships

In the Measurement Information Model of ISO 15939, the link between the two major parts illustrated in Figure 2 (that is, the Metrology-related bottom part of measurement and the Analysis-related upper part of quantification) are not explicitly described. ISO 15939 makes the assumption that this link exists and that it is complete on its own. In practice, the issue might be more complex, in particular in domains where measurement or quantification (or both) are not yet mature. In practice, there is no guarantee that what can be measured
adequately at the level of base and derived quantities indeed represents the concepts and relationships that the analysis part of the Measurement Information Model attempts to quantify. An example of this is the maintainability characteristic in ISO 9126, which is not strictly limited to the software entity itself, but is implicitly related to the entity ‘effort required to maintain such software at a later time’.

6 A Productivity Model and the ISO 15939 Measurement Information Model

This section illustrates the differences between the metrology and non-metrology-related part of the ISO 15939 Measurement Information Model by looking at the differences between a productivity ratio and a productivity model. It also illustrates how statistical techniques propose standard reference models to facilitate the analysis of relationships embedded in software production and estimation models.

6.1 A productivity model is more than a productivity ratio

There are major differences between a productivity ratio and a productivity model. A productivity ratio is strictly related to the metrology-related part of the ISO 15939 Measurement Information Model: it is strictly defined as composed of two base measures (Output over Input):

- This productivity ratio is based strictly on the measurement, from a metrology perspective, of the respective single attribute of the corresponding two distinct entities representing the output and the input.
- This productivity ratio is strictly descriptive and limited to what is being measured.

If we move now from a productivity ratio to a productivity model, a number of additional elements are added, since the purpose of this analysis model becomes:

- the analysis of relationships across many entities (e.g. many projects) that have been measured, and often
- an estimation of what would happen should this production process (which has been measured through its output and input) be used again to estimate the next project.

To explore this, let us look at Figure 4, which illustrates a production process in its simplest form: The Input is on the left, the Process is in the middle, and the Output is on the right. It is to be noted that, while the productivity ratio has two explicit dimensions (Input and Output) that are explicitly present in a productivity ratio, the production model has another (implicit) dimension as well, that is, the production process itself.

![Figure 4: A (Software) Production Process](image)

The objective is now to quantify this production model, normally using a productivity model (more commonly referred to as an estimation model in software engineering), rather than measuring it. This quantitative representation of the production process is typically built by:

- collecting the base measures of the production process over a number of completed projects, for example: Input = Effort (in work-hours or work-days) and Output = Functional Size of the completed software (in Function Points)
- quantitatively modeling the relationships across these two variables, where effort is the dependent variable, and functional size is the independent variable.

Statistical techniques can be considered as standard reference models for modeling these relationships, each with a distinct mathematical representation (and corresponding strengths, constraints, and limitations). The productivity model, in its simplest form with a single dependent variable \(x\), could be expressed as: \(x = f(y)\), where

- the dependent variable \(x\) would be in work-hours, and
- the independent variable \(y\) would represent the size of the software (in Function Points).

Quantification of the productivity ratio is the outcome of the measurement of two entities (that is, Inputs and Outputs), while the meaning of the division of these two numbers represents something different, that is, the performance (in the sense of productivity) of a third entity, the process itself. This means that the measurement of the productivity of the process is derived not from a direct measurement of the process, but from an indirect measurement of two other entities (the inputs and outputs of the process).

The Analysis model is expected to consider a number of distinct dimensions and combine them, in
some manner, into a single number. This corresponds to the definition in ISO 15939, which is that a model is an algorithm combining one or more base and/or derived measures, along with their associated decision criteria.

Finally, it must be observed that, while a derived measure gives a combination of units (e.g. Function Points per work-hour), the productivity model produces as output a single quantity with its corresponding single unit of the dependent variable. Here, the output of the F(x) analysis model is strictly in work-hours (Effort), even though many additional independent variables would have been taken into account in more comprehensive productivity models (i.e. estimation models).

6.2 A productivity model built with a linear regression.

This subsection presents an example of a productivity model built with the linear regression technique from a set of completed projects. This linear regression statistical function can be considered as the algorithm of a standard reference model.

A- The standard reference model – linear regression

A productivity model built using the linear regression technique is presented in Figure 5. The quantitative representation from the linear regression statistical technique is of the following form: the dependent variable of Effort is a function of the independent variable of functional Size, that is: Effort = f(Functional Size). In the linear regression model, this equation takes the following quantitative form:

Effort = a x Functional Size + b

In practical terms, in this equation from the linear regression,

- a represents the slope of the linear regression line, and
- b represents the point at the origin (that is, when the independent variable = 0)

Stated differently, the slope a represents the increase of 1 unit of effort for an increase of 1 unit of functional size.

In terms of measurement units, this equation then corresponds to Effort (in work-days) =

\[ a \times \text{(work-hours/Function Point)} \times \text{Functional size (in Function Points)} + b \times \text{(work-days at the origin when functional size = 0)}. \]

When the mathematical expression is worked out with its measurement units, then the end result is, indeed, in work-days. Therefore, both the left- and right-hand sides of the equation have the same measurement unit, which is ‘work-days’.

Finally, there is also a practical interpretation of this production model:

- slope a represents the variable cost of the production process (that is, the variation due to an increase in functional size), and
- the b value at the origin represents the fixed cost (that is, the portion of effort that is independent of the variation of the functional size).

For such a standard reference model (i.e. the linear regression model), a number of evaluation criteria of such statistical models are available in the literature, such as the coefficient of determination (R²), the magnitude of the relative error (MRE), the mean magnitude of the relative error (MMRE), and the predictive quality of the model (PRED). There are various interpretations of the values of these evaluation criteria: in the software engineering literature, an estimation model is generally considered good when the MRE is within +/-25% for 75% of the observations, or when PRED(0.25) = 0.75.

B- The organizational reference context

A specific productivity model built within an organization with the set of completed projects from this organization becomes the organizational reference context. Such a productivity model can then be used for estimating the next project for this organization. This productivity model will

- provide a specific estimate that would be directly on the linear regression line, as well as
- provide various elements of information on the quality characteristics of this model (such as its R², MMRE, etc.), which could be used as additional elements to make a decision on whether or not the selected estimate for the specific project would be above the regression line (i.e. more costly), on the regression line, or below the regression line (i.e. less costly).

With respect to the ISO 15939 Measurement Information Model,

- the regression model would correspond to the Analysis Model,
• the specific outcome of the regression model would be the Indicator, and
• the set of information from the specific productivity model built by this organization would correspond to the Interpretation context, while the standard statistical technique of linear regression, which forms the basis for the organizational reference context, would also be part of the Interpretation context.

7 Summary
Practitioners and managers can have much more confidence in measurement results based on international standards of measurement (i.e. metrology) than on the judgment-based quantification of implicit relationships across a mix of entities and attributes without due consideration of admissible mathematical operations on numbers of different scale types. This paper has analyzed the Measurement Information Model in ISO 15939 and clarified what in it refers to the classical metrology field, and what refers to quantitative analysis of relationships. The example presented has illustrated the quantification of relationships modeled by a productivity model built from a well-known and well-mastered statistical analysis technique such as regression. However, some of the quantitative models, such as the quality models built using ISO 9126, will provide numbers which typically depend on a specific selection among a (potentially large) number of alternatives, the assignment of a percentage to each contributing alternative, based on the opinion of one person (or a group of people), and comparison of the contributing alternatives with their distinct threshold values that are often defined by opinion as well. Future work is required to investigate, for instance, if these numbers, when combined adequately, consider the admissible mathematical operations and treatment of the corresponding measurement units, and whether or not the outcomes of such models are indeed meaningful numbers.

References