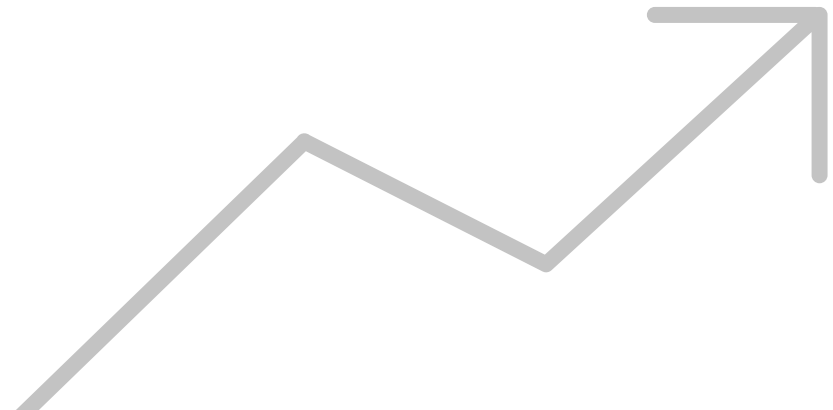


Quality Framework for Statistical Algorithms

ML2022 Meeting, Star Wars Day 2022

Wesley Yung, Siu-Ming Tam, Bart Buelens, Hugh Chipman, **Florian Dumpert**, Gabriele Ascari, Fabiana Rocci, Joep Burger and InKyung Choi



Starting Point

Existing Quality Frameworks for Institutes, Processes and Products

To mention some (that have been “represented” in the quality group) :

- Australian Bureau of Statistics ' Data Quality Framework
- United Nations ' National Quality Assurance Framework
- Eurostat's European Statistics Code of Practice
- Statistics Canada's Quality Assurance Framework

Starting Point

Eurostat's European Statistics Code of Practice

Aspects on processes:

- Sound Methodology
- Appropriate Statistical Procedures
- Non-excessive Burden on Respondents
- Cost Effectiveness

Aspects on statistical output:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability
- Accessibility and Clarity

Our Running Example:

An employee of a statistical office transferring descriptions of the business of an enterprise to the corresponding NACE code, i.e. to the correct statistical classification of economic activities.

“Our firm gathers berries and nuts and sells them to our customers.”



02.30 Gathering of wild growing non-wood products

“Our firm is engaged in growing berries and nuts and sells them to our customers.”



01.25 Growing of other tree and bush fruits and nuts

Key Question

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Key Question

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Aspects on processes:

- Sound Methodology
- Appropriate Statistical Procedures
- Non-excessive Burden on Respondents
- Cost Effectiveness

Aspects on statistical output:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability
- Accessibility and Clarity

Key Question

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Sound Methodology

- Sound Methodology

Statistical authorities maintain and develop cooperation with the scientific community to improve methodology, the effectiveness of the methods implemented and to promote better tools when feasible.

- Non-excessive Burden on Respondents

(Principle 7.7 of the European statistics Code of Practice)

- Cost Effectiveness

Aspects on statistical output:

- Relevance

- Accuracy and Reliability

- Timeliness and Punctuality

- Coherence and Comparability

- Accessibility and Clarity

Key Question

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Appropriate Statistical Procedures

- Sound Methodology
- Appropriate Statistical Procedures
- Non-excessive Burden on Respondents
- Cost Effectiveness

Statistical processes are routinely monitored and revised as required.

(Principle 8.3 of the European statistics Code of Practice)

Aspects on statistical output:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability
- Accessibility and Clarity

Key Question

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Non-excessive Burden on Respondents

- Sound Methodology

Administrative and other data sources are used whenever possible to avoid duplicating requests for data.

- Appropriate Statistical Procedures

Statistical authorities promote measures that enable the linking of data sources in order to minimise response burden.

- Non-excessive Burden on Respondents

(Principles 9.4 and 9.6 of the European statistics Code of Practice)

- Cost Effectiveness

Aspects on statistical output:

- Relevance

Administrative and other data sources are used whenever possible to avoid duplicating requests for data.

- Accuracy and Reliability

Administrative and other data sources are used whenever possible to avoid duplicating requests for data.

- Timeliness and Periodicity

Administrative and other data sources are used whenever possible to avoid duplicating requests for data.

- Coherence and Comparability

Administrative and other data sources are used whenever possible to avoid duplicating requests for data.

- Accessibility and Clarity

Key Question

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Cost Effectiveness

- Sound Methodology

The productivity potential of information and communication technology is being optimised for the statistical processes.

- Appropriate Statistical Processes

- Non-excessive Burden on Respondents

- Cost Effectiveness

(Principle 10.2 of the European statistics Code of Practice)

Aspects on statistical output:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability
- Accessibility and Clarity

Intermediate Result No. 1

(a)

We have to consider statistical algorithms in all steps of the GSBPM.

(b)

We have to do that a meaningful manner (in particular in terms of quality).

A Meaningful Manner?

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Aspects on processes:

- Sound Methodology
- Appropriate Statistical Procedures
- Non-excessive Burden on Respondents
- Cost Effectiveness

Aspects on statistical output:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability
- Accessibility and Clarity

A Meaningful Manner?

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Aspects on statistical input: Accuracy and Reliability

- Sound Methodology
- Appropriate Statistical Procedures
- Non-excessive Burden on Respondents
- Cost Effectiveness

European Statistics accurately and reliably portray reality.

(Principle 12 of the European statistics Code of Practice)

Aspects on statistical output:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability
- Accessibility and Clarity

A Meaningful Manner?

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Timeliness and Punctuality

- Sound Methodology
- Appropriate Statistical Procedures
- Non-excessive Burden on Respondents
- Cost Effectiveness

European Statistics are released in a timely and punctual manner.

(Principle 13 of the European statistics Code of Practice)

Aspects on statistical output:

- Relevance
- Accuracy and Reliability
- Timeliness and Punctuality
- Coherence and Comparability
- Accessibility and Clarity

A Meaningful Manner?

How are quality aspects affected if we use statistical algorithms in some steps of the GSBPM?

Accessibility and Clarity

- Sound Methodology

Statistics and the corresponding metadata are presented, and archived, in a form that facilitates proper interpretation and meaningful comparisons.

- Non-excessive Burden on Respondents

(Principle 15.1 of the European statistics Code of Practice)

- Cost Effectiveness

Aspects on statistical output:

- Relevance

Statistics and the corresponding metadata are presented, and archived, in a form that facilitates proper interpretation and meaningful comparisons.

- Accuracy and Reliability
- Timeliness and Punctuality

- Coherence and Comparability

- Accessibility and Clarity

Intermediate Result No. 2

When we want to use statistical algorithms (and we should, see Intermediate Result No. 1), the quality requirements on processes and the statistical outputs have to be fulfilled .

(Nothing new, right?)

Quality Framework for Statistical Algorithms (QF4SA)

Or: How to ensure this?

- Developed in 2019/2020 in WP 2 of the UNECE HLMOS Machine Learning Project
- Work package lead: Wesley Yung (STATCAN)
- Published in SJIAOS and by UNECE

1) DOI 10.3233/SJT210875

2) <https://unece.org/statistics/publications/machine-learning-official-statistics>

Quality Framework for Statistical Algorithms (QF4SA)

Or: How to ensure this?

5 quality dimensions

- Accuracy
- Explainability
- Reproducibility
- Timeliness
- Cost effectiveness

Quality Framework for Statistical Algorithms (QF4SA)

Accuracy

The accuracy of statistical information refers to the **degree to which it correctly describes the phenomena it was designed to measure**; i.e., it is the closeness of computations or estimates to the exact or true values that the statistics were intended to measure.

Quality Framework for Statistical Algorithms (QF4SA)

Accuracy

Aspects/Metrics:

- Depending on the task (e.g., MSE, MAE, R^2 ; Acc, F1.;)
- Includes point estimation and its uncertainty
- Often estimated by using resampling methods
- Should be communicated

But what about our
employee's accuracy?

Quality Framework for Statistical Algorithms (QF4SA)

Explainability

Defined as the **ability to understand the logic underpinning** the algorithm used in prediction or analysis, as well as the resulting outputs, greatly assisted by depicting **the relationship between the input and output variables** and providing the necessary information on the methodology underpinning the algorithm.

Quality Framework for Statistical Algorithms (QF4SA)

Explainability

Aspects/Metrics:

- Degree to which a human can understand how a prediction is made from a statistical or an ML algorithm using its input features.
- Difficult to measure by a number (between 0 and 1)
- Different assisting approaches (LIME, PDP, ...)

But what about our
employee's explainability ?

Quality Framework for Statistical Algorithms (QF4SA)

Reproducibility

At the basic level defined as the **ability to replicate results using the same data and algorithm** originally used.

(Higher Levels are possible.)

Quality Framework for Statistical Algorithms (QF4SA)

Reproducibility

Aspects/Metrics:

- Reproducibility is a minimum necessary condition for a finding to be believable and informative
- For official statistics to be trusted, they need to be reproducible – trust is the currency of official statistics

Quality Framework for Statistical Algorithms (QF4SA)

Timeliness

Defined by the **time involved in producing a result from conceptualization to algorithm building, processing and production** . A distinction should be made between timeliness in development and production.

Quality Framework for Statistical Algorithms (QF4SA)

Timeliness

Aspects/Metrics:

- With the increased use of large datasets, the speed at which ML algorithms can be trained and run could lead to significant improvement in timeliness (e.g., for coding)
- Once developed (!), ML algorithms can process vast amounts of data in a short time

Quality Framework for Statistical Algorithms (QF4SA)

Cost effectiveness

Defined as the degree to which the results are effective in relation to their cost, i.e., **accuracy (*or timeliness, or ...*) per unit cost.**

Quality Framework for Statistical Algorithms (QF4SA)

Cost effectiveness

Aspects/Metrics:

- Fixed costs, such as infrastructure and staff training
- Ongoing costs, such as production costs

Conclusion

- QF4SA does not replace existing quality frameworks but complements them with a special view on statistical algorithms
- QF4SA works for “traditional ” algorithms as well as for machine learning
- QF4SA can and should be used for comparing different algorithms (perhaps including jobs that are done manually)

Outlook

- There may be other more appropriate definitions for these dimensions.
- There may be other important dimensions of quality not mentioned here.
- There may be results or experiences not known by the authors of QF4SA.

Outlook



Quality Aspects of Machine Learning -
Official Statistics between Specific Quality Requirements and Methodological Innovation

CALL FOR PAPERS

On the one hand, machine learning methods offer new possibilities for analysis and insight – not only but also against the background of new types of data (sources), some of which are immense in scope – and thus enable higher-quality statistics and can contribute to efficiency improvements in the processes of official statistics. On the other hand, it has not yet been generally clarified whether and how the typically prediction-oriented, non-model-based approach of machine learning methods can be reconciled with the special quality requirements and framework conditions of official statistics. Keywords here are, for example:

- legal requirements regarding sample sizes, characteristics to be surveyed, or publication deadlines to be met;
- requirements for the accuracy of estimates to be achieved;
- frequent use of complex sample designs;
- requirements for transparency of processes and methods.

With the European Code of Practice, official statistics in Europe are subject to special quality requirements. One of these requirements is that the statistical processes for collecting, processing and disseminating statistics fully comply with international standards and guidelines and at the same time reflect the current state of scientific research.

In order to bring together the already built-up empirical knowledge from the practice of official statistics with the findings of science, and thus to promote and intensify the exchange between researchers from the field of machine learning and official statisticians working on applications of modern machine learning methods, Statistics Network Bavaria intends to organize a scientific workshop from 6th to 8th September 2022 in Munich. The workshop will be organized for Statistics Network Bavaria by the Bavarian State Statistical Office, the Institute of Statistics of the Ludwig-Maximilians-University Munich and the ifo Institute Munich with the cooperation of the Federal Statistical Office of Germany.

In order to ensure the richest and most profitable exchange possible, to concretize and, if necessary, expand the abstract questions mentioned above as examples, and thus to promote future joint (research) work, the organizers are looking forward to relevant contributions from official statistics as well as from academia in the form of presentations and discussions within the framework of the workshop.

The organizers are in discussion with the editors of the journal *ASTA Wirtschafts- und Sozialstatistisches Archiv* regarding a special issue on the topic. Papers presented at this workshop are eligible for publication in this special issue, subject to successful completion of the review process.

Please send proposals by means of an abstract (1 to 2 pages) to statistiknetzwerk@statistik.bayern.de by May 22nd, 2022 at the latest. Abstracts should be submitted in German or English. Since we expect guests from non-German speaking countries, it is expected that slides will be provided at least in English (we are happy to additionally provide slides in German). Presentations can be given in German or English.

Contact persons:

- Thomas Augustin (Institute of Statistics, LMU Munich)
- Sebastian Wichert (ifo Institute – Leibniz Institute for Economic Research at the University of Munich)
- Nina Storfinger (Bavarian State Statistical Office, Statistics Network Bavaria)
- Florian Dumpert (Federal Statistical Office of Germany, AI and Big Data Unit)

For more information on the Statistics Network Bavaria, please visit: <https://www.statistiknetzwerk.bayern.de>

Workshop on Quality Aspects of Machine Learning Official Statistics between Specific Quality Requirements and Methodological Innovation September 6th to 8th, 2022 in Munich (Germany)

https://www.statistiknetzwerk.bayern.de/mam/themen/workshops/2022_cfp_en.pdf

Contact

Statistisches Bundesamt
Federal Statistical Office
65180 Wiesbaden
Germany

www.destatis.de

www.destatis.de/kontakt

Contact Person
Florian Dumpert
florian.dumpert@destatis.de
Phone +49 611 75-3887

