

Use of the correlation matrix approach to define the life detection techniques in a sample curation facility

John Robert Brucato (1), **Andrea Meneghin** (1), Sara Russell (2), Caroline Smith (2), Petra Rettberg (3), Allan Bennet (4), Tom Pottage (4), Aurore Hutzler (5) and the EURO-CARES Team

(1) INAF - Astrophysical Observatory of Arcetri, Firenze, Italy (meneghin@arcteri.astro.it), (2) Natural History Museum, London, UK, (3) Deutsches Zentrum für Luft und Raumfahrt, Cologne, Germany, (4) Public Health England, Salisbury, UK, (5) Naturhistorisches Museum Wien, Vienna, Austria

Abstract

EURO-CARES (European Curation of Astromaterials Returned from Exploration of Space) was a three year (2015-2017), multinational project, funded under the European Commission's Horizon 2020 research programme to develop a roadmap for a European Extra-terrestrial Sample Curation Facility (ESCF). If the samples are brought back to Earth from bodies where there is the possibility of presence of extant or extinct life, there are a wide number of proposed approaches on the techniques to use in order to investigate the presence of biosignatures: [3], [4], [5], etc. All the studies lead to a proposed list of techniques suitable for life detection along with details about the field of application, their efficiency and limits. What is missing is a critical approach able to make a comparison between the techniques in terms of effectiveness, to find a prioritizing ranking. In this paper a quality engineering tool approach, the correlation matrix, was used to support the choice of the techniques for life detection, [1], [2]. The challenge was to analyze and evaluate every technique. To do it, a wide panel of expert was involved. Experts in the following scientific and technological field composed the team: process engineering, mechanical engineering, biology, astrobiology, chemistry. The paper shows how, using a logical flow of analysis, it was possible to identify the critical issues and to highlight the priorities.

1. Introduction

The major drivers we took into account were to define which techniques are really important and which can be considered as optional, rationalize the activity flow inside the curation and provide a support for the design choices of the curation.

Starting from this idea, we focused on the building of a correlation matrix where to correlate the biosignatures with the available techniques. It is known that a number of techniques can detect each biosignature and, at the same time, each technique can be applied for a number of biosignatures. Using the correlation matrix method it is possible to summarize all this information at a glance. It is also possible to give an extra-value to the matrix, trying to be more critical: the idea is not only to determine the correlations between the biosignatures and the techniques, but also to define how strong is each correlation.

2. The correlation matrix

The correlation matrix (Figure 1) shows the correlation between biosignatures and the life detection techniques. According to the matrix approach, the biosignatures were organized per area (morphological, chemical, biochemical, isotopic analysis, and mineralogical), an importance value was given to each techniques, in a range from 1 to 4, and a correlation value was defined, in an exponential range from 0 to 9: 0 if no correlation exists, 1 (low correlation) if the technique is no specific for the biosignature but still usable and/or with medium/low resolution, 3 (medium correlation) if the technique is suitable for the biosignature, although not specific, and/or with medium resolution and 9 (high correlation) if the technique is very specific technique for the biosignature, with high resolution. An extra value was given to disentangle destructive and non-destructive techniques, (1 if the technique is destructive, 1.1 if partially destructive, 1.2 if partially destructive/non-destructive, 1.3 if non-destructive). The numerical results obtained from the correlation matrix are the biosignature occurrence (number of times that the each biosignature is

detected by a different techniques), the techniques occurrences (the number of biosignatures that can be detected by a single techniques), the technique mean value (the technique mean correlation with the detected biosignatures) and finally the technique importance rating calculated, for each column (technique), as the sum of the products of the biosignatures importance, the correlation value and the non-destructive/destructive coefficient.

TECHNIQUES	IMPORTANCE	BIOSIGNATURES OCCURRENCES																										
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Size of single cell - Size of targets	4																											
Numbers of single bacteria - Number of targets	4																											
Population size (Colonies)	4																											
Chemical	4																											
- Cellular composition	4																											
- Organic molecules	4																											
- DNA, RNA	4																											
Biochemical	4																											
- Organic pigments	4																											
- Protein	4																											
Isotopic analysis	3																											
- Isotopes, isotopologues	3																											
- Isotopomers	2																											
Elemental analysis	4																											
- Structure, Mineralogy	4																											
Mineralogical	4																											
Optical microscopy - SEM	4	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
Electron microscopy - SEM	4	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
GC-MS	4																											
LC-MS	4																											
MALDI-TOF	4																											
Fluorescence microscopy	4																											
Raman spectroscopy	4																											
High Performance liquid Chromatography (HPLC)	4																											
Polymerase Chain Reaction (PCR)	4																											
Enzyme-linked immunosorbent assays (ELISA)	4																											
Fluorescent in-situ hybridization (FISH)	4																											
Sequencing	4																											
Chromatography	4																											
Protein microarray / Marker Chip	4																											
SIMS	4																											
Isotope Ratio Mass Spectrometry (IRMS)	4																											
13C-NMR	4																											
SEM-EDX	4																											
XRF	4																											
X-Ray CT	4																											
XRD	4																											
Electron microscopy - TEM	4																											
NMR	4																											
FTIR	4																											
Marker chip with antibody	4																											
Capillary Electrophoresis (CE)	4																											
MC-ICP-MS	4																											
TECHNIQUES OCCURRENCES		4	3	2	4	4	3	2	4	4	3	2	4	4	3	2	4	4	3	2	4	4	3	2	4	4	3	2
TECHNIQUES MEAN VALUE		7.5	9.0	7.0	9.0	6.0	7.0	9.0	6.0	7.0	9.0	6.0	7.0	9.0	6.0	7.0	9.0	6.0	7.0	9.0	6.0	7.0	9.0	6.0	7.0	9.0	6.0	7.0
TECHNIQUES IMPORTANCE RATINGS		102	81	72	96	64	84	86	72	36	72	40	36	31	27	18	44	52	36	36	42	8	43	12	12	9	1	1
NON DESTRUCTIVE/DESTRUCTIVE COEFFICIENT		1.3	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
TECHNIQUES IMPORTANCE RATING (*ND/D)		138	99	81	72	96	83	101	48	36	72	36	72	40	36	31	27	23	48	52	47	47	55	10	56	12	12	9

Figure 1: biosignatures/techniques correlation matrix.

3. Main results

Starting from the numerical results it is possible to make some observations: the initial list included 27 different techniques but the he number of high-correlation techniques are 21; there are 8 techniques able to detect 2 or more biosignatures; if only high-correlation techniques are considered, the minimum

number of techniques needed to detect all the biosignatures is 9. This number decreases to 7, if also the medium correlation techniques are considered; if only the high-correlation techniques are considered, the minimum number of techniques needed to solve all the high-importance (given value 4) biosignatures is 7, and this number decreases to 6 if also the medium correlation (given value 3) techniques are considered.

4. Summary and Conclusions

The proposed correlation matrix technique is a powerful tool able to convert a subjective approach to an objective one, helping to rationalize a problem from the boundaries definition to the final solution. The matrix allows to select the most important techniques. This leads us to define the procedures to be performed inside the ESCF, which are strictly related to the techniques. Starting from the obtained results it is possible to facilitate the design choices: choosing a technique allows a better evaluation of curation dimensions (depending on the size and position of the instrument, etc.) and layout (depending on its position, the compatibility with other instruments, the need of ancillary systems, etc.).

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 640190.

References

- [1] Akao Y., Development History of Quality Function Deployment - The Customer Driven Approach to Quality Planning and Deployment, Asian Productivity Organization. p. 339. ISBN 9283311213, 1994.
- [2] Cohen L. et al., Quality Function Deployment and Six Sigma: A QFD Handbook, Prentice Hall, ISBN: 9780137035441, 2009.
- [3] Kminek G. et al., Report of the workshop for life detection in samples from Mars, 2014.
- [4] Race M. et al., Assessment of Planetary Protection Requirements for Mars Sample Return Missions, National Research Council, 2009.
- [5] Wilson A. et al., Exobiology in the Solar System & The Search for Life on Mars, ESA, ISBN 9290925205, 1999.