https://dx.doi.org/10.22161/ijaers.5.4.18 ISSN: 2349-6495(P) | 2456-1908(O)

Three-Parameter Logistic Model (ML3): A Bibliometrics Analysis

Edson Luis Leopardi Medeiros[1], Guilherme de Sá[1], Pedro Gabriel Ambrosio[1], Leopoldo Pedro Guimarães Filho[1], Vilson Menegon Bristot[1], Kristian Madeira[1], Cristina Keiko Yamaguchi[1,2], Fernanda Cristina Silva Ferreira[2], Stéfano Frizzo Stefenon[3]

[1]University of Extremo Sul Catarinense (UNESC), Criciúma, Brazil
[2]University of Planalto Catarinense (UNIPLAC), Lages, Brazil
[3]Santa Catarina State University (UDESC), Joinville, Brazil

***Abstract*— *In the current scenario of Production Engineering, where analysis tools are applied like indicators of production, control and optimization of productive processes or even in the application of advanced analytical methods to support decision making, the Three-parameter Logistic Model (ML3) is one of the most promising, belonging to the Item Response Theory (IRT). This model provides ways of representing the relation between the probabilities of an individual giving a certain response to an item. Also it can identify latent features and parameters of the items, in the area of study knowledge. The objective of this research was to verify the scientific productions found in the electronic databases selected around the Three Parameters Logistic Model. A bibliometrics was carried in the year of 2012 until the last publication of the month of July of 2017. The bases that were investigated are Web of science, Scopus, SciELO and Google Scholar. The authors found twelve articles about the subject. The year of 2016 was the most productive (33,3%), articles with three authors were the most frequent (41,7%), SAGE Journals Educational and Psychological Measurement was the journal with the highest number of publications (25.0%), the USA was the most productive country (41.7%) and Yorkville University was the most profitable university (16.7%). The conclusion is that although the ML3 is relevant for the development of several areas of Production Engineering in relation to IRT, it still little used and exploited in the production of scientific knowledge, revealing a potential area for the development of new researches.*

Keywords*— *Item Response Theory (IRT), Three-parameter Logistic Model, Bibliometrics, Production Engineering.

I. INTRODUCTION

According to Tezza and Bornia [1], identifying problems and proposing improvements depends on high-level surveys in the data collection so that the method of analysis must be appropriate to the scenario to give credibility to the information. That is why data collection and the analysis method are crucial to propose a solution to a given problem, or to suggest an improvement in some process.

Among the instruments used in Production Engineering to build production indicators, to plan, control and optimize production processes or even to apply advanced analytical methods to support decision making, one can highlight the measurement scales. They enable the verification of divergences and the comparison of data through relations and thus propose solutions. In this scenario is the Item Response Theory (IRT).

IRT uses measurement scales to propose methods, in order to analyze, verify and find viable solutions for data analysis, and later for a more effective decision-making process. In Brazil, in addition to being used in school evaluations and psychometric tests, v has also been used in areas such as services, total quality management and evaluation of intangibles in organizations [1].

IRT is a set of mathematical models that represents the probability of an individual giving a certain response to an item as a function of item parameters and the respondent's abilities. This link is always expressed in such a way that the greater the skill, the greater the likelihood of success in the item [2].

The various models proposed in scientific studies and articles depend fundamentally on three factors: (a) universe of the item (dichotomous or non-dichotomous); (b) number of populations involved; and (c) the amount of latent traits being measured.

IRT models follow two assumptions, which refer to the characteristics of the items. The first is associated with one-dimensionality, that is, the group of items must measure the same variable. Although there is a skill set behind any behavioral performance, it is assumed that a single skill is being measured to satisfy the one-dimensionality assumption. The second one refers to the local independence of items. This means that a response to an item has no influence on the responses given to

https://dx.doi.org/10.22161/ijaers.5.4.18 *ISSN: 2349-6495(P) | 2456-1908(O)*

other items. Taking this assumption as true, one infers that the sequence of responses of the subject to a series of items will be the product of the probabilities of each item individually. There is no way to demonstrate these two assumptions, they are simply accepted or not [3].

From some models proposed by the IRT, the Three-parameter Logistic Model (ML3) is one of the most used. In the year 1952 Frederic Lord develops a two-parameter IRT model supported by the normal cumulative distribution, but with some uses of the two-parameter model the same author understands the imperative to incorporate a parameter that deals with the adversity of the casual hit. Thus, the three-parameter model emerged.

Although researches on the ML3 do exist, there is no knowledge about statistics regarding scientific production on this subject. This work developed a bibliometric study of scientific production on the ML3 from January 2012 to June 2017. The electronic databases used were Web of Science, Science Direct, Scopus, SciELO and Google Scholar.

The objectives of this work were: (a) quantifying articles about ML3 per year; (b) identifying the most productive authors on the subject; (c) verifying the number of authors by scientific production; (d) verifying the scientific journals and journals with the highest number of articles published; (e) quantifying the most productive countries; (f) verifying the institutions with the largest number of articles published; and (g) observing the H index of the most productive authors.

II. BIBLIOMETRIC STUDY

According to Araújo and Alvarenga [4], bibliometrics is a research tool with indicators that aim to portray the behavior and development of an area of human knowledge. As stated by Andrade [5], these indicators evaluate authors' productivity, quantification of citations, and frequency of the appearance of keywords, among others.

In general, with the evolution of information and advances in scientific production, together with the expansion of scientific databases, areas such as bibliometrics, scientometrics and webometrics are increasingly important [4]. Although these areas have their specificities, they have in common the interest in disseminating knowledge, identifying the main tendencies, structural and quantitative characteristics, regarding scientific articles published in congresses and scientific journals [6].

As stated by Vanti [7], among the several applications of bibliometrics, the most relevant were: (a) identify trends and the growth of knowledge in an area; (b) identify journals at the core of a discipline; (c) measure magazine coverage; (c) predict publication trends; (d) measure the degree and patterns of collaboration between authors; and (e) measuring the growth of certain areas and the emergence of new subjects.

In addition, bibliometrics has several laws and conceptions that employ statistical and mathematical processes dictating foundations of research and ordering in the scientific analyzes on the information universe. Lotka, Bradford and Zipf's Laws [8] are among the most employed laws.

According to Silva et al. [6], Lotka's Law, also known as the Law of Inverse Square started in 1926, where one was found that most published scientific research is produced by a small number of authors, and conversely, most of the researchers do not have the same income as the great researchers, but their research is equivalent to the sum of the publications.

According to Guedes and Borschevier [8], Bradford's Law allows establishing the degree of relevance of journals in a defined area of knowledge, in which the periodicals that publish the greater number of articles on a certain subject form a nucleus of periodicals, presumably of higher quality or relevance to the area.

Zipf's Law or minimum effort consists in measuring the frequency of the emergence of keywords in various texts. For this purpose, a list is organized by terms of a given area. This law is further divided into two others, where the first law states that a word that appears repeatedly in the same text exposes the subject of the publication. The second law reports that in a certain publication, several words of low repetitiveness found in the text have the same frequency [7].

III. ITEM RESPONSE THEORY

The Item Response Theory indicates models for the traits suggesting ways of representing the relationship between the probability of a individual to give an answer to an item and its latent traits or abilities, in the area of knowledge to be evaluated or ascertained, which cannot be directly observed [9].

The applications of techniques derived from IRT are in many areas of knowledge, such as educational, medical, psychosocial, marketing, services and total quality management.

One of the advantages of IRT is the possibility of making comparisons between abilities of individuals of different populations when they are submitted to tests having common items, and the comparison of individuals of the same population undergoing totally different tests. This is possible because IRT has as its central elements the items, instead of the global evaluation [10].

The item response theory indicates aspects still little explored by Production Engineering, although it portrays advances in aspects of control and management of

https://dx.doi.org/10.22161/ijaers.5.4.18 ISSN: 2349-6495(P) | 2456-1908(O)

resources and processes, such as the "power to position individuals or processes of different groups on a common scale, even though these have responded to different items, allowing the identification of opportunities for improvement or even benchmarking; allowing a more accurate evaluation of properties' items and their results and, consequently, allow greater application of statistical techniques; understand the psychometric properties of the instruments; the possibility of developing more efficient indicators to evaluate individual differences in processes, practices, systems or individuals; and greater robustness of the results" [1].

According to Araujo, Andrade and Bortolloti [10], IRT models lack the type of item and type of response method, being they cumulative or not. The fundamental discrepancy between the two models lies in the connection of the assertive response probability given by an individual to an item in relation to its characteristics in cumulative models. The increase in probability is related to the growth of the latent trait and the item's parameter characteristics. For non-cumulative models, in some cases the probability is, for example, related to the function of the distance between the parameters and the item in the scale and do not exactly depends exclusively on the parameters and the latent trait [11]. IRT offers mathematical models for latent traits, providing ways of representing the relationship between the probability of an individual to give a certain response to an item, its latent trait and item's characteristics (parameters), in the area of study knowledge [10,11].

Of the models proposed by the IRT, the one-dimensional three-parameter logistic model (ML3), the most used nowadays, is given by mathematical formula (1):

$$P\left(Uij = 1 \middle| \theta_j\right) = c_i + \left(1 - c_i\right) \cdot \frac{1}{1 + e^{-D a_i \left(\theta_j - b_i\right)}} \quad (1)$$

Source: ANDRADE et al, [11].

As $i = 1, 2, \cdots, I$, e $j = 1, 2, \cdots, n$, where:

U_{ij}: is a dichotomous variable that assumes the values 1, when the individual *j* correctly answers the item *i*, or 0 when the individual *j* does not respond correctly to item *i*.

θ_j: represents the ability (latent trait) of the j-th individual.

$P(U_{ij} = 1|\theta_j)$: is the probability that an individual *j* with skill θj correctly responds to item *i* and is called the Item Response Function.

b_i: is the parameter of difficulty (or position) of item *i*, measured in the same skill scale.

a_i: is the discrimination (or inclination) parameter of item *i*, with a value proportional to the slope of the Item Characteristic Curve - CCI at point *bi*.

c_i: is the parameter of the item that represents the probability of individuals with low ability to correctly respond to item *i* (often referred to as casual hit probability).

D: is a scale factor, constant and equal to 1. The value 1.7 is used when the logistic function is expected to deliver results similar to the normal warhead function [11].

IV. METHODOLOGICAL PROCEDURES

The present study was carried out based on research in four electronic databases: Scopus, Web of Science, Science Direct and SciELO, besides Google Scholar. The accessions were carried out in the Laboratory of the Center of Studies in Production Engineering (NEEP) of UNESC (Universidade do Extremo Sul Catarinense) from Criciúma, Santa Catarina state. The period of verification of the articles was from January 2012 to June 2017.

The keywords used to perform the search in the databases and in Google Scholar were combined with the Boolean operator "AND" and resulted in the following search expression: "Engineering AND Three Parameter Logistic Model". The research included only scientific papers published in journals, excluding, therefore, citations and patents, chapters of books and publications in annals of scientific meetings and congresses. The selected job information has been exported to Microsoft Excel 2013 software, for further statistic treatment.

V. RESULTS AND DISCUSSIONS

The search strategy used in this research ("Engineering AND Three Parameter Logistic Model") resulted in 10 (ten) articles from Science Direct, 1 (one) article from Web of Science, 4 (four) articles from Scopus, 0 (zero) article from SciELO and 99 (ninety-nine) articles from Google Scholar, totaling 114 articles, being that 12 (doze) addressed the Three-parameter Logistic Model. Table 1 presents data on the publications of the 12 scientific papers analyzed.

Table.1: Number of articles published on ML3 in the period from January 2012 to June 2017.

	2012	2013	2014	2015	2016	2017*	Total
Articles	1	2	2	2	4	1	12

* Values up to June.

Source: Search data, 2017.

Observing Table 1, the publications related to the proposed subject have been maintaining the subject present over the years in the scientific scenario, with publications throughout the analyzed period, offering an average of two publications per semester, being the year of 2016 with the largest number of publications (n = 4). The articles published in the year 2016 approach the application of the Three-parameter Logistic Model (ML3) in educational and psychological measurements. IRT models help in this purpose by analyzing items and producing a standardized scale. An elaboration of the items must be carried out by a professional who has knowledge of the subject and must follow techniques for the elaboration of items [3].

In addition to the number of articles published per year, the study has also investigated which were the most productive authors. Table 2 presents the most productive authors with their publications analyzed annually.

Table.2: Number of articles published on ML3 in the period from January 2012 to June 2017.

Authors	2012	2013	2014	2015	2016	2017*	Total
Audrey J. Leroux	0	1	0	0	0	0	1
Christy Brown	0	0	0	1	0	0	1
Kpolovie, P.J	0	0	1	0	0	0	1
Kyung Yong Kim	0	0	0	0	0	1	1
Louis Tay	0	0	0	0	1	0	1
Rita de Cássia Correa	0	0	0	0	1	0	1
Sandip Sinharay	0	0	0	1	0	0	1
Shana Moothedath	0	0	0	0	1	0	1
Suttida Rakkapao	0	0	0	0	1	0	1
Tianheng Wang	0	1	0	0	0	0	1
Ting-Wei Chiu	1	0	0	0	0	0	1
Zeki Kaya	0	0	1	0	0	0	1

* Values up to June.

Source: Search data, 2017.

.

Table 2 shows a balance in the number of publications among authors who published in the field of Three-parameter Logistic Model (ML3) in the period from 2012 to 2017. Each author has published only one article on the topic. Although no exponent has been identified for productivity in the ML3 area, this type of analysis is very important and needs to be addressed in a bibliometric study in accordance with Lotka's Law [7]. The productivity of scientific authors is verified more formally by calculating the H-index, that is, this index measures the impact of the scientists on their peers based on their most cited articles. In this sense, although a balance was observed in the number of articles published among the authors, the most prominent H-index found were the ones from scientists Sandip Sinharay (H = 16) and Louis Tay (H = 15).

Another aspect investigated in the bibliometrics held was the quantity of authors by scientific production. The curiosity was based on the assumption that it would be necessary at least two researchers for the accomplishment of a bibliometrics, being one with experience in one area of knowledge and another, a beginner, oriented by the first one.

Table 3 - Number of authors per article on ML3 from January 2012 to June 2017.

Number of authors	2012	2013	2014	2015	2016	2017*	Total
One author	0	0	0	1	0	0	1
Two authors	1	0	2	0	0	1	4
Three authors	0	1	0	1	3	0	5
Four authors	0	1	0	0	0	0	1
Five authors	0	0	0	0	1	0	1

* Values up to June.

Source: Search data, 2017.

According to the values of Table 3, that the publications appearing in greater quantity with 3 authors was verified, in sequence with two authors, and finishing 1.4 and five authors. Thus, the articles with 3 authors

https://dx.doi.org/10.22161/ijaers.5.4.18 ISSN: 2349-6495(P) | 2456-1908(O)

corresponded to 41.67% of the total of publications, articles with two authors represented 33.33% of the total articles published, followed by articles with 1.4 and 5 authors with 8.33% that adding represent 25% of the total articles published in the period.

Another aspect investigated in the present bibliometrics was the journals with greater amount of publications on ML3. Table 4 presents the results of this analysis.

Table.4: Journals with the largest number of ML3 publications - From January, 2012 to June, 2017.

Journals	2012	2013	2014	2015	2016	2017*	Total
Applied Measurement in Education	0	0	0	0	0	1	1
Biomedical Optics Express	0	1	0	0	0	0	1
European Journal of Statistics and Probability	0	0	1	0	0	0	1
Perspectives in Education	0	0	0	0	1	0	1
Physical Review Physics Education Research	0	0	0	0	1	0	1
Journal of Education and Research in Accounting	0	0	0	0	1	0	1
SAGE Journals Applied Psychological Measurement	0	0	0	1	0	0	1
SAGE Journals Educational and Psychological Measurement	1	1	0	0	1	0	3
SAGE Journals Journal of Educational and Behavioral Statistics	0	0	0	1	0	0	1
Turkish Online Journal of Distance Education-TOJDE	0	0	1	0	0	0	1

* Values up to June.

Source: Search data, 2017.

According to Table 4, data show that only the *Educational and Psychological Measurement* journal had the number of publications above the average, with three articles published in the years of 2012, 2013 and 2016, representing 25% of the publications. *Educational and Psychological Measurement* is a publication of SAGE JOURNALS. Sara Miller McCune founded SAGE in 1965 in order to Support the dissemination of usable knowledge and to educate a global community. SAGE is one of the leading publishers of innovative and quality content, covering a wide range of subject areas. Corroborating the quality of the publisher, Table 4 shows two more magazines of its catalog to *Applied Psychological Measurement* and the *Journal of Educational and Behavioral Statistics*, with two other articles published. Thus, the sum of the other magazines or newspapers represents 58.3% of the total of publications.

Regarding the quality of the journals, one can see a variation of the impact factor of 0,885 (*SAGE Journals Educational and Psychological Measurement* e *SAGE Journals Applied Psychological Measurement*) to 3,337 (*Biomedical Optics Express*). The impact factor of four journals was not found (Perspectives in Education, Physical Review Physics Education Research, Turkish Online Journal of Distance Education-TOJDE and European Journal of Statistics and Probability).

Another aspect investigated in the present study was the most productive countries. The results of this investigation are shown in Table 5.

Table.5: Most productive countries in ML3 publications from January, 2012 to June, 2017.

Countries	2012	2013	2014	2015	2016	2017*	Total
Brazil	0	0	0	0	1	0	1
Canada	1	0	0	0	1	0	2
USA	0	2	0	2	0	1	5
India	0	0	0	0	1	0	1
Nigeria	0	0	1	0	0	0	1
Tailand	0	0	0	0	1	0	1
Turkey	0	0	1	0	0	0	1

* Values up to June.

Source: Search data, 2017.

https://dx.doi.org/10.22161/ijaers.5.4.18 ISSN: 2349-6495(P) | 2456-1908(O)

Table 5 clearly shows that the United States of America is the country with the highest number of productions, with five publications, representing 41.67% of the amount of total research published in the analyzed period. Canada was the second most productive country with two articles and 16.67% of the total articles published in the period evaluated. The other countries, Brazil, India, Nigeria, Thailand and Turkey with one article each, accounted for 41.67% of the total articles published in the period evaluated. Thus, the most productive countries were the USA and Canada. Renowned universities in the areas of research and structure suitable for scientific production can justify this data.

The most productive research institutions were also investigated. Table 6 presents a summary of the results obtained.

Table.6: Most productive institutions in ML3 publications from January 2012 to June 2017.

Institution	2012	2013	2014	2015	2016	2017*	Total
Clemson University	0	0	0	1	0	0	1
Gazi Üniversitesi	0	0	1	0	0	0	1
Indian Institute of Technology	0	0	0	0	1	0	1
Pacific Metrics Corporation	0	0	0	1	0	0	1
Prince of Songkla University	0	0	0	0	1	0	1
The University of Texas at Austin	0	1	0	0	0	0	1
Universidade Federal de Santa Catarina	0	0	0	0	1	0	1
University of Connecticut	0	1	0	0	0	0	1
University of Iowa	0	0	0	0	0	1	1
University of Port Harcourt	0	0	1	0	0	0	1
Yorkville University	1	0	0	0	1	0	2

* Values up to June.

Source: Search data, 2017.

One could verify, from Table 6, that the Yorkville University from Canada was the institution that most stood out, to include two articles, representing 16.67% of the total articles published on the subject investigated in the period assessed.

Finally, the key words that were highlighted in the scientific articles were evaluated. In this analysis, the 12 articles were collected regardless of the year of publication. Table 7 presents the results.

Table 7. Number of keywords in ML3 publications from January 2012 to June 2017.

Keywords	Total
Item	7
Model	6
Theory	6
Response	6
Three-parameter	2
Tomography	2
Logistic	2
Optical	2
Test	2
Other words	49

Source: Search data, 2017.

With respect to the keywords, the most frequent ones were: item (n = 7), model (n = 6), theory (n = 6), response (n = 6), three-parameter (n = 2), tomography (n = 2), logistic (n = 2), optical (n = 2) and test (n = 2). The other words (n = 49) appeared only once each. One can realize from the number of keywords found that the Item Response Theory (IRT), more specifically the Three-parameter Logistic Model (ML3), presents a broad spectrum of applications.

VI. CONCLUSION

This article has as main objective to present a bibliometric study of articles published between the years of 2012 and June 2017 on the Three-parameter Logistic Model (ML3). The articles were classified and evaluated by year, author, and number of authors by scientific production, journals and newspapers with the highest amount of articles published, most productive countries and institutions with the highest number of articles published on the subject.

One can notice that largest number of publications occurred in the year 2016; The USA had the largest number of scientific articles published in the area; the most profitable institution was the Canadian Yorkville University; and no author exponent in the area of Three-parameter Logistic Model (ML3) was found by the number of published articles. However, the authors Sandip Sinharay and Louis Tay stood out by H-index. Another important finding was that, although journal impact factors ranged from 0.885 to 3.337, this indicator was not found in four journals.

With this research, one can conclude that there is a gap to be explored by researchers in the area of Production Engineering regarding the Item Response Theory, which are the Three-parameter Logistic Models. This subject is very important for Production Engineering, as stated by Tezza and Bornia (2009), and can assist in the control and management of resources and processes, such as the positioning of individuals or processes of different groups on a common scale, even if these individuals have responded to different items, allowing identification of opportunities for improvement or even benchmarking. This subject can also allow a more accurate evaluation of the items' properties and their results, allowing greater precision in the application of statistical techniques. It can also assist in the adequate understanding of instruments' psychometric properties; enable the development of more efficient indicators to evaluate individual differences in processes, practices, systems or individuals; and to conclude, it can provide more robust results.

This study presents as limitation the small number of articles on the theme, which did not allow a general mapping on the topic. The authors suggest that further research on the subject be included in papers published in annals of events, monographs, theses and dissertations, books, book chapters and patents.

This research was very important to reveal a gap in the field of research in the area of Production Engineering, which must be occupied in order to contribute to the growth of this field.

REFERENCES

[1] Tezza, R., Bornia, A. C. (2009). Teoria da Resposta ao Item: vantagens e oportunidades para a engenharia de produção. XXIX Encontro Nacional de Engenharia de Produção–ENEGEP.

[2] Gomes, L. S., Carvalho, P. C. P. (2014). A Teoria de Resposta ao Item na avaliação em larga escala: um estudo sobre o Exame Nacional de Acesso do Mestrado Profissional em Matemática em Rede Nacional – PROFMAT 2012. 2014. 79 f. Dissertação (Mestrado) - Curso de Mestrado Profissional em Matemática, Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro.

[3] Pasquali, L., Primi, R. (2003). Fundamentos da teoria da resposta ao item: TRI. Avaliação Psicológica, v. 2, n. 2, p. 99-110.

[4] Araújo, R. F., Alvarenga, L. A. (2011). Bibliometria na pesquisa científica da pós-graduação brasileira de 1987 a 2007. Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf., v. 16, n. 31, p. 51-70.

[5] Andrade, F. S., Jung, C. F. (2012). Análise bibliométrica da produção científica de pesquisadores e referências de um periódico da engenharia de produção. 2012. 75 f. Dissertação (Mestrado) - Curso de Programa de Pós-graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul, Rio Grande do Sul.

[6] Silva, J. M. P. et al. (2012). O estado-da-arte da literatura em economia e gestão da inovação e tecnologia: um estudo bibliométrico.

[7] Vanti, Nadia Aurora Peres (2002). Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. Ciência da informação, v. 31, n. 2, p. 152-162.

[8] Guedes, V. L. S., Borschiver, S. (2005). Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica. Encontro Nacional de Ciência da Informação, v. 6, p. 1-18.

[9] Spenassato, D., Kinas, P. G. (2009). Teoria da Resposta ao Item (TRI): Estimação Bayesiana da habilidade de indivíduos. Vetor, Rio Grande, v. 19, n. 2, p.74-84.

[10] Araujo, E. A. C., Andrade, D. F., Bortolotti, S. L. V. (2009). Teoria da Resposta ao Item. Rev Esc Enferm Usp, Florianópolis, v. 43, n. spe, p.1000-1008.

[11] Andrade, D. F., Valle, R. C., Tavares, H. R. (2000). Introdução à teoria da resposta ao ítem: conceitos e aplicações. SINAPE. Retrieved from http://homes.ufam.edu.br/jcardoso/LivroTRI.pdf.