

Optimizing Search Results Using Taguchi Design of Experiments: Case Study of the Distribution of COVID-19 Names in Vietnam

Ton Nguyen Trong Hien
Department of Business Administration
Van Lang University
Ho Chi Minh city, Vietnam
hien.tnt@vlu.edu.vn

Abstract—Taguchi design of experiments is a method used in engineering, especially in the field of new product development, for optimizing production processes and improving product quality. The Taguchi method is widely used in the field of industrial engineering. The purpose of this study is to investigate the applicability of the Taguchi method for other tasks, such as optimizing search results. Using the Taguchi method, we assess the popularity of three names for COVID-19 that have been used by scientists in research related to COVID-19 in the context of Vietnam. The results show the positive value of using the Taguchi method to optimize search results. The keyword COVID-19 is found to be the most-used keyword. The results also show that disease prevention, spread, infection, transmission, treatments, preparation, and responses to the disease are particularly interesting research topics in Vietnam at present.

Index Terms—COVID-19, Design of Experiments, Taguchi, Vietnam.

I. INTRODUCTION

ASSOCIATED with the development of mankind is the manufacturing industry. Manufacturers prefer to have the best product quality at a reasonable cost, and a process of quality management is needed to achieve quality control. It is crucial that consumers are satisfied with their purchases; however, if the partial targets deviate from the target value, customers will feel dissatisfied and the cost of loss will also increase. A system of statistical tools, including the Taguchi method, has been developed for quality control.

Because quality control is of concern in engineering, most of the previous research has been based on the application of the Taguchi method in engineering fields. The main research question of this study is whether the Taguchi method can be applied to other fields. To answer this question, we consider the interesting case study of searching for the most common name of COVID-19 in Vietnam.

II. LITERATURE REVIEW

The Taguchi design of experiments is a statistical method developed by its inventor, Genichi Taguchi, who refined his ideas in the 1950s. His method was virtually unknown outside of Japan until the early 1980s [1]. The challenge of modern industry is mainly focused on achieving high quality, saving costs, and increasing production efficiency. The Taguchi method is a useful tool, widely used in manufacturing, to optimize designs for performance and quality, limit product defects, and minimize production costs with limited experiments [2]–[5]. To confirm this, we obtained 4,020 re-

sults from the Scopus database, nearly 90% of which are research studies published in the field of engineering and technology. The results are shown in Fig. 1.

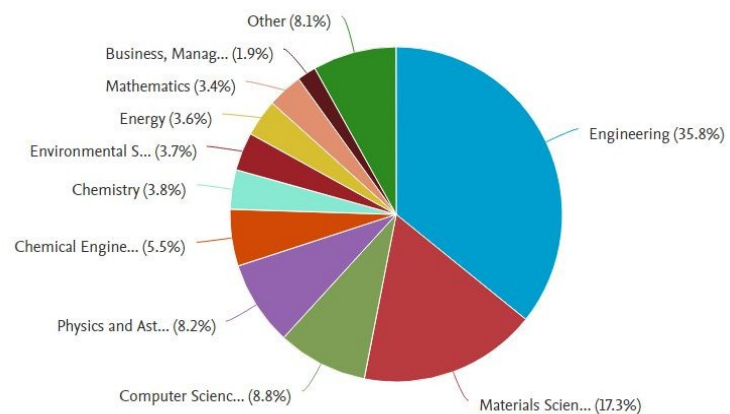


Fig. 1 The pie diagram shows the distribution among various fields of the use of the Taguchi method (Source: Scopus, extraction date: 1 July 2021)

Since the Taguchi method is based on statistics rather than modeling, it is very popular in engineering. In discussing the Taguchi method, one immediately thinks of technical systems development [6]. When it comes to quality management, it was W. Edwards Deming who first recognized the importance of reversing quality control from inspection to process control. However, Taguchi enhanced quality control when he devised the control tool at the design stage. Broadly speaking, the Taguchi method should not be limited to manufacturing engineering alone as it can be applied to any process where controlling the early design phase is important.

The COVID-19 epidemic, the name given by the World Health Organization (WHO), also known as coronavirus 2019 and SARS-CoV-2, is an infectious disease caused by a newly discovered coronavirus [7]–[9]. The disease originated in Wuhan, China, spread throughout many provinces and cities in that country, and not long afterward spread widely to many other countries. At the present time, in mid-July 2021, there have been millions of cases. Consequently, with the ongoing COVID-19 outbreak, healthcare systems across the world have been pushed to the brink. Furthermore, it costs low- and middle-income countries nearly US\$52 billion (US\$8.60/person) every four weeks to deliver an effective

healthcare response to COVID-19 [10]. The epidemic is posing unprecedented challenges and creating enormous difficulties in all aspects of life. This has encouraged scientists to conduct research. According to the Dimensions database, about 4% of the world's research output was devoted to coronavirus in 2020, and a sharp increase in the number of articles submitted to scientific journals on coronavirus can be expected [11]. Vietnam is no exception. At present, Vietnam is experiencing a fourth fast-spreading wave of infections, and there have been several reputable studies related to COVID-19 in the Vietnamese context.

The name for the COVID-19 epidemic has not been agreed upon by scientists [12]. As explained by the WHO, viruses and diseases often have different names. For example, HIV is the virus that causes AIDS. People often know the name of a disease, but not the name of the virus that causes it. There are different procedures and purposes for naming viruses and diseases. Viruses are named based on their genetic makeup to facilitate the development of diagnostic tests, vaccines, and drugs. Diseases are named so that prevention, transmission, severity, treatments, preparations, and responses can be discussed [13]. Therefore, it is important to utilize the prevailing name to assess current research and to identify research gaps that need to be filled.

Conducting a systematic review in any field of scholarly research is paramount. In the research design phase, it can be difficult to deliver a meticulous summary of all available primary research in a succinct and readable form in response to a research question. There are many ways to locate search terms. Therefore, the purpose of this study is to investigate the use of the Taguchi method to evaluate search terms. To illustrate this approach, this paper presents an experiment on optimizing search results given the distribution of COVID-19 names in Vietnam.

III. PROPOSED WORK

Design of experiments is a statistical approach to investigate a system or process to assess the relationship between multiple input variables (factors) and key output variables (responses) so that the influence can be found. Taguchi experimental design can be used at the point of greatest leverage to reduce design costs by speeding up the design process, as the Taguchi method is designed to solve this problem with a small number of trials [5], [14]. It is not necessary to perform all test cases for all scenarios. Results show that the Taguchi method gives similar results and saves 88.9% on test runs compared with other methods [5].

TABLE I. KEYWORD AND TRIAL EXPERIMENTS

Term Experiment		SARS- COV-2	COVID-19	Coronavirus 2019
1	Vietnam	AND	AND	AND
2		AND	AND	NOT
3		AND	NOT	AND
4		AND	NOT	NOT
5		NOT	AND	AND
6		NOT	AND	NOT
7		NOT	NOT	AND

The details of the seven trials are shown in Table 1. In the trials, we specified the following search terms: "SARS-COV-2," "COVID-19," "Coronavirus 2019," and the fixed keyword, "Vietnam."

Scenarios for a full factorial design (Table 1) are as follows:

- 1) "Vietnam" AND "SARS-COV-2" AND "COVID-19" AND "Coronavirus 2019."
- 2) "Vietnam" AND "SARS-COV-2" AND "COVID-19" NOT "Coronavirus 2019."
- 3) "Vietnam" AND "SARS-COV-2" NOT "COVID-19" AND "Coronavirus 2019."
- 4) "Vietnam" AND "SARS-COV-2" NOT "COVID-19" NOT "Coronavirus 2019."
- 5) "Vietnam" NOT "SARS-COV-2" AND "COVID-19" AND "Coronavirus 2019."
- 6) "Vietnam" NOT "SARS-COV-2" AND "COVID-19" NOT "Coronavirus 2019."
- 7) "Vietnam" NOT "SARS-COV-2" NOT "COVID-19" AND "Coronavirus 2019."

Data on COVID-19 research questions were collected from a variety of sources in selected journal databases. In choosing databases for the experiment, we selected three reputable, multidisciplinary publishers/databases as data sources: Taylor and Francis, Emerald, and ScienceDirect. Specifically, Taylor and Francis is a UK publisher providing access to more than 2,700 high-quality, multidisciplinary journals in the humanities and social sciences, natural sciences and technology, engineering, and medicine and healthcare. Emerald, dating back to 1967, is a long-established UK publisher of journals and books in the fields of management, business, education, healthcare, and engineering. ScienceDirect is a website that provides access to a large database of multidisciplinary scientific and medical publications by the Dutch publishing house, Elsevier. The database stores more than 18 million articles from more than 4,000 academic journals and 30,000 e-books. Thus, the selection of data sources provides the assurance of large size, inclusiveness, and academic reliability. The search time was 16h00' on 15 July 2021 (GMT +7). Statistical results are shown in Table 2. The results were processed with Minitab 18 software.

IV. RESULTS AND DISCUSSION

Evaluating keyword changes means looking at how adding or subtracting keywords affects the results. Fig. 2 shows a downward-sloping line for the keyword "COVID-19" from the first level (AND) to the second level (NOT). A

TABLE II. DESIGN OF EXPERIMENT RESULTS

Database Experiment	Taylor and Francis	Emerald	ScienceDirect
1	0	0	0
2	0	0	1
3	0	0	0
4	0	0	2
5	0	0	0
6	16	8	23
7	1	0	0

line that is not horizontal (not parallel to the x-axis) indicates that there is an effect. By comparing the slopes of the lines, it is possible to compare the relative magnitude of the factor effects. The larger the slope, the greater the difference in results when changing one or more keywords. This difference can be clearly seen if we remove the design parameter keywords, “SARS-COV-2” or “Coronavirus 2019,” in a search string. As can be seen in Fig. 2, the steeper lines for the two keywords, “SARS-COV-2” and “Coronavirus 2019,” are parallel to each other. The parallelism shows non-interaction, which means they have a similar result or consequence. However, the longer line representing the keyword “Coronavirus 2019” indicates that if this word is removed, the settings will be optimal. The figure also shows that the keyword “COVID-19” has a positive signal when it exists in the search string; in other words, it makes one of the most significant contributions.

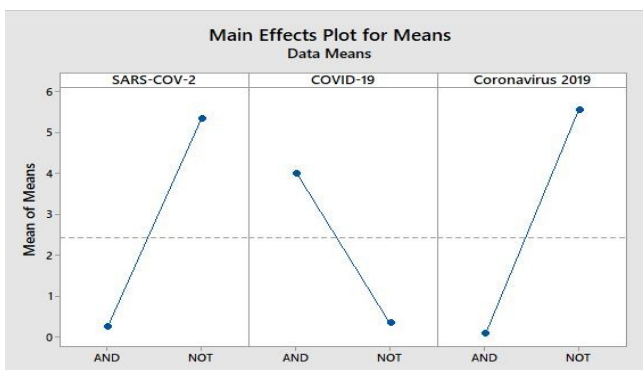


Fig. 2 Main effects plot for the case study.

In the search results, the phrase “COVID-19” generates the most responses across the databases. According to the WHO definition, the word “COVID-19” implies disease. Problems currently discussed in the context of Vietnam are disease prevention, disease spread with transmission, and illness severity. It can be suggested that pandemics prompt the government to find ways to treat, prepare for, and respond to diseases.

V. CONCLUSION AND FUTURE WORK

The Taguchi method is a technique to optimize the product design phase in manufacturing, in particular, and engineering, in general. This article has demonstrated the use of the Taguchi method for optimizing search array strings as a way of applying the technique to another field. The results

of the case study show that the keyword “COVID-19” returns the optimal syntax. Hopefully, this can answer the question of whether the Taguchi method can be applied to other fields.

The fight against COVID is not over yet. The case study shows that the response to prevent and control the COVID-19 epidemic remains a matter of concern. In the future, we suggest empirical research to test the effectiveness of the Taguchi method for different nonengineering fields, such as education, healthcare, etc.

REFERENCES

- [1] J. Antony, D. Perry, C. Wang, and M. Kumar, “An application of Taguchi method of experimental design for new product design and development process,” *Assem. Autom.*, vol. 26, no. 1, pp. 18-24, Jan. 2006.
- [2] A. Jerbi, A. Ammar, M. Krid, and B. Salah, “Performance optimization of a flexible manufacturing system using simulation: The Taguchi method versus OptQuest,” *Simulation*, vol. 95, no. 11, pp. 1085-1096, Jan. 2019.
- [3] I. Choudhury and M. El-Baradie, “Surface roughness prediction in the turning of high-strength steel by factorial design of experiments,” *J. Mater. Process. Technol.*, vol. 67, no. 1-3, pp. 55-61, 1997.
- [4] J. Antony and F. Jiju Antony, “Teaching the Taguchi method to industrial engineers,” *Work Study*, vol. 50, no. 4, pp. 141-149, July 2001.
- [5] M. Islam and A. Pramanik, “Comparison of design of experiments via traditional and Taguchi method,” *J. Adv. Manuf. Syst.*, vol. 15, no. 3, pp. 151-160, Sept. 2016.
- [6] V. Silva, *Statistical approaches with emphasis on design of experiments applied to chemical processes*. IntechOpen, 2018, p. 142.
- [7] B. Hu, H. Guo, P. Zhou, and Z. Shi, “Characteristics of SARS-CoV-2 and COVID-19,” *Nat. Rev. Microbiol.*, vol. 19, no. 3, pp. 141-154, Oct. 2020.
- [8] T. Singhal, “A review of coronavirus disease-2019 (COVID-19),” *Indian J. Pediatr.* vol. 87, no. 4, pp. 281-286, Apr. 2020.
- [9] Y. Wu, C. Chen, and Y. Chan, “The outbreak of COVID-19: An overview,” *J. Chin. Med. Assoc.*, vol. 83, no. 3, pp. 217-220, Mar. 2020.
- [10] A. Kaye et al., “Economic impact of COVID-19 pandemic on healthcare facilities and systems: International perspectives,” *Best Pract. Res. Clin. Anaesthesiol.*, vol. 35, no. 3, pp. 293-306, Nov. 2021.
- [11] H. Else, “How a torrent of COVID science changed research publishing — in seven charts,” *Nature*, vol. 588, no. 7839, 553, Dec. 2020.
- [12] S. Jiang et al., “A distinct name is needed for the new coronavirus,” *The Lancet*, vol. 395, no. 10228, p. 949, Feb. 2020.
- [13] “Naming the coronavirus disease (COVID-19) and the virus that causes it,” WHO, 2021. [Online]. Available: [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it).
- [14] N. Fei, N. Mehat, and S. Kamaruddin, “Practical applications of Taguchi method for optimization of processing parameters for plastic injection moulding: A retrospective review,” *ISRN Ind. Eng.*, vol. 2013, 462174, June 2013.