Hindawi
Complexity
Volume 2021, Article ID 6351836, 8 pages
https://doi.org/10.1155/2021/6351836

Research Article

A Statistical Approach to Model the *H*-Index Based on the Total Number of Citations and the Duration from the Publishing of the First Article

Mohammad Reza Mahmoudi,[1] **Marzieh Rahmati**,[2] **Zulkefli Mansor**,[3] **Amirhosein Mosavi**,[4,5] **and Shahab S. Band**[6,7]

[1]*Department of Statistics, Faculty of Science, Fasa University, Fasa, Fars, Iran*
[2]*Department of Computer Engineering, Yazd University, Yazd, Iran*
[3]*Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia*
[4]*Environmental Quality, Atmospheric Science and Climate Change Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[5]*Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[6]*Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam*
[7]*Future Technology Research Center, College of Future, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan*

Correspondence should be addressed to Amirhosein Mosavi; amirhosein.mosavi@tdtu.edu.vn and Shahab S. Band; shahabsband@yahoo.com

Received 6 May 2020; Revised 24 December 2020; Accepted 19 February 2021; Published 1 March 2021

Academic Editor: Diego R. Amancio

The productivity of researchers and the impact of the work they do are a preoccupation of universities, research funding agencies, and sometimes even researchers themselves. The *h*-index (h) is the most popular of different metrics to measure these activities. This research deals with presenting a practical approach to model the *h*-index based on the total number of citations (N_C) and the duration from the publishing of the first article (D_1). To determine the effect of every factor (N_C and D_1) on h, we applied a set of simple nonlinear regression. The results indicated that both N_C and D_1 had a significant effect on h ($p < 0.001$). The determination of coefficient for these equations to estimate the *h*-index was 93.4% and 39.8%, respectively, which verified that the model based on N_C had a better fit. Then, to record the simultaneous effects of N_C and D_1 on h, multiple nonlinear regression was applied. The results indicated that N_C and D_1 had a significant effect on h ($p < 0.001$). Also, the determination of coefficient for this equation to estimate h was 93.6%. Finally, to model and estimate the *h*-index, as a function of N_C and D_1, multiple nonlinear quartile regression was used. The goodness of the fitted model was also assessed.

1. Introduction

The productivity of researchers and the impact of the work they do are a preoccupation of universities, research funding agencies, and sometimes even researchers themselves. Several metrics have been used to measure these including journal impact factors, citation counts, and publication rates. At present, however, the *h*-index is the most popular of these metrics [1–4]. Hirsch's definition of the index is that $h = m$ if m of a researcher's p papers have at least m citations each and each of the other papers has no more than m citations. As a guide, Hirsch [1] suggested that a "successful" scientist would have $h = 20$ after 20 years of work, whereas outstanding and "truly unique" individuals would have $h = 40$ and $h = 60$, respectively, after 20 years of work. Subsequent work has shown that this is too great generalisation, if only because the *h*-index is highly discipline-specific and depends on circumstance, the comprehensiveness of the literature databases is used to calculate the index and many others [5, 6]. For example, very eminent mathematicians often have $h < 10$ and some Nobel laureates also have very small h-indices [7]. The inevitable inference is an individual's h-index should be considered in the context of these factors and of the distribution of h for a given number of papers and citations appropriate to the individual researcher. Some researchers introduced alternative versions of the h [8]. Generally, all of the given indices consider the number of citations received by articles. Recently, scientist researchers have studied and developed theoretical models to estimate and model these indices based on other indicators, for example, based on N_C [1], the total number of publications, T [9], and the total number of publications with a minimum of one citation, T_1 [10]; based on N_C and T [4, 11–15]; and based on N_C, T_1, and the total number of citations for the 1 most cited papers, C_1 [16]. Librarians are particularly interested in using good tools to predict future individual scientific achievements. To solve this problem, Hirsch [17] indicated that the h-index acts significantly better than other alternatives including N_C, T, and mean citations per paper to forecast future scientific achievement. It has been shown that the h-index is better than other alternatives to predict productivity.

This research deals with a statistical approach to model the h-index based on N_C and D_1. Simple and multiple nonlinear regressions and multiple nonlinear quartile regression were applied to estimate and predict the h-index based on N_C and D_1. The results were also compared to the results of simple and multiple linear regressions (SLR and MLR) models as common methods.

2. Methodology

This section is devoted to discussing the details of data collection, samples, and statistical techniques that have been applied to analyze the dataset.

2.1. Data Collection. The dataset of this work contains the information of articles for 29470 Iranian scientists that have been indexed in Google Scholar.

2.2. Data Analysis. Statistics and data mining are popular approaches to extract knowledge from the dataset. These approaches contain different data analysis techniques such as descriptive statistics [18–22], regression models [23–29], time series models [30–43], and clustering analysis [44]. In this research, the data gathered from Google Scholar were fed and analyzed using the *SPSS 25* and *R 3.3.2* software.

The descriptive statistics of research variables contained minimum, maximum, mean, standard deviation, and quartiles are reported in Table 1. As Table 1 indicates, the means of h, N_C, and D_1 for Iranian scientists are 5.74, 248.78, and 7.98, respectively. Also, the value of h for at least 25%, 50%, and 75% of them is at most 2 ($Q_1 = 2$), 4 ($Q_2 = 4$), and 7 ($Q_3 = 7$), respectively.

To determine the effect of every factor (N_C and D_1) on h, we applied a set of simple nonlinear regression. Also, to investigate the simultaneous effects of N_C and D_1 on h, multiple nonlinear regressions were applied. Finally, we divided the observations into 4 groups as follows: the first group: observations with $h \le 2$; the second group: observations with $2 < h \le 4$; the third group: observations with $4 < h \le 7$; the fourth group: observations with $h > 7$. Then, to model and estimate the h based on N_C and D_1, the multiple nonlinear quartile regression (MNLQR) was used. The goodness of applied models was also evaluated by the coefficient of determination (R^2) and comparing actual values with predicted values. The accuracies of the models were investigated using the five-fold cross-validation. In other words, the dataset was divided into five parts. In each step, four parts were considered as training data and the other part was considered as testing data. The models were trained using training data and the trained models were applied on testing data. Finally, the discrepancy between the predicted h and the true h were evaluated using different measures such as R^2, root mean square error (RMSE), and mean absolute error (MAE).

2.2.1. Simple Nonlinear Regression. To model a quantitative response variable Y based on a predictor variable X, simple nonlinear regression (SNLR) model is a powerful technique. The general equation of SNLR is presented by

$$Y = \beta_0 + \beta_1 X^{\beta_2} + \varepsilon, \tag{1}$$

where β_0, β_1, and β_2 are model parameters and ε is the random error. The estimated equation of the SLR model is given by

$$\widehat{Y} = b_0 + b_1 X^{b_2}, \tag{2}$$

where b_0, b_1, b_2, and $\widehat{Y}$ are estimations of β_0, β_1, β_2, and Y, respectively.

2.2.2. Multiple Nonlinear Regression. To model a quantitative response variable Y based on predictor variables $X_1, \ldots, X_k$, multiple nonlinear regression (MNLR) is a powerful technique. The general equation of MNLR with two predictors X_1 and X_2 is presented by

$$Y = \beta_0 + \beta_1 X J_1^{\beta_2} + \beta_3 X_2^{\beta_4} + \beta_5 X_1^{\beta_6} X_2^{\beta_7} + \varepsilon, \tag{3}$$

where $\beta_0, \ldots, \beta_7$ are model parameters and ε is the random error. The estimated equation of the MNLR model is also given by

$$\widehat{Y} = b_0 + b_1 X_1^{b_2} + b_3 {X_2}^{b_4} + b_5 X_1^{b_6} X_1^{b_7}, \tag{4}$$

where $b_0, \ldots, b_7$ are estimations of $\beta_0, \ldots, \beta_7$ and $\widehat{Y}$ is the estimated value of Y.

2.2.3. Multiple Nonlinear Quartile Regression. In multiple nonlinear quartile regression (MNLQR), first, the quartiles of response variable have been computed. Then, based on the values of quartiles, the observations were categorized into 4

Table 1: Descriptive statistics of variables in the dataset.

	Minimum	Maximum	Mean	Std. deviation	Quartile First (Q_1)	Second (Q_2)	Third (Q_3)
H	1	84	5.74	5.79	2.00	4.00	7.00
N_C	1	37570	248.78	828.84	15.00	58.00	200.00
D_1	1	42	7.98	4.59	5.00	7.00	10.00

Table 2: The results of SNLR models to study the effect of N_C and D_1 on h.

Factor	Source	Sum of squares	df	Mean squares	F	R^2	p
N_C	Regression	1892965	2	946482.5	429143.66	0.934	<0.001
	Residual	64992.09	29468	2.205514			
	Uncorrected total	1957957	29470				
	Corrected total	987069.3	29469				
D_1	Regression	1364231	2	682115.4	33854.95	0.398	<0.001
	Residual	593726.3	29468	20.14817			
	Uncorrected total	1957957	29470				
	Corrected total	987069.3	29469				

Table 3: The parameter estimates of SNLR models for N_C and D_1.

Factor	Parameter	Estimate	Std. error	95% confidence interval Lower bound	Upper bound	p
N_C	b_1	0.600	0.003	0.595	0.606	<0.001
	b_2	0.476	0.001	0.474	0.477	<0.001
D_1	b_1	0.667	0.014	0.640	0.694	<0.001
	b_2	1.041	0.008	1.025	1.057	<0.001

distinct categories. Finally, a separate MNLR was fit for each category.

3. Results

The SNLR results to predict the separate effects of every factor (N_C and D_1) on h are given in the first section. Section 2 is in regard to the MNLR results to investigate the simultaneous effects of N_C and D_1 on h. Section 3 reports the MNLQR results to model the effects of factors on h, in each quartile.

3.1. SNLR Results. This part is to study the impact of each factor (N_C and D_1) on h. In this research, h was the response variable. Also the variables N_C and D_1 were continuous predictors. Tables 2 and 3 summarize the results of SNLR models for the variables N_C and D_1. As Table 2 indicates, N_C and D_1 factors had a significant effect on h ($p < 0.001$). Figure 1 also shows the plot of the fitted curve with data.

Table 3 shows the parameter estimates of SNLR models for N_C and D_1, respectively. Based on the results of Table 3, we can estimate h as a function of N_C and D_1, by

$$\widehat{h}_{N_C} = 0.600 N_C^{0.476}, \tag{5}$$

$$\widehat{h}_{D_1} = 0.667 D_1^{1.041}, \tag{6}$$

respectively. Also, the R^2 values for these equations to estimate h were 93.4% and 39.8%, respectively.

Table 4 shows the results of SLR as a comparative method. Based on the results of Table 4, we can estimate h as a function of N_C and D_1, by

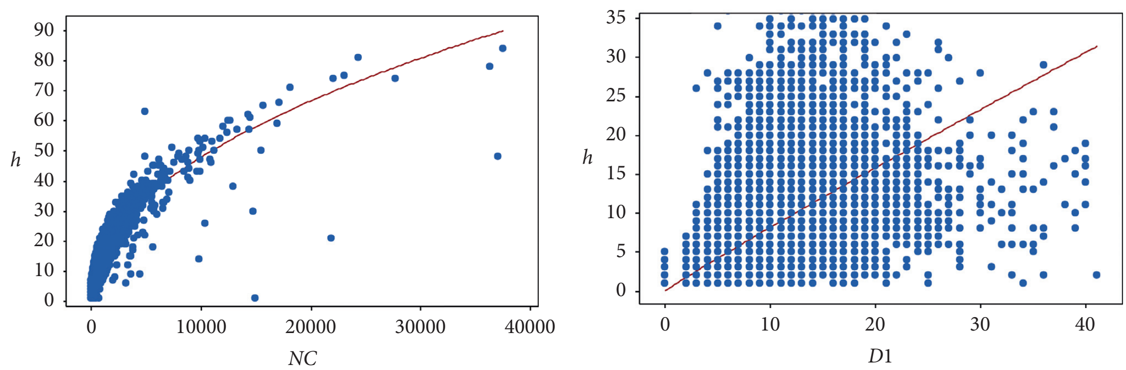
$$\widehat{h}_{N_C} = 4.429 + 0.005 N_C, \tag{7}$$

$$\widehat{h}_{D_1} = 0.172 + 0.797 D_1, \tag{8}$$

respectively. Also, the R^2 values for these equations to estimate of h were 56.9% and 24.9%, respectively. As it can be observed, the SNLR method acts better than the SLR method.

Table 5 summarizes the results of five-fold cross-validation. The results confirm that the SNLR method acts better than the SLR method.

Figure 2 and Table 6 show the plot of actual values versus predicted values and the correlations between them. As can be seen, the SNLR model based on N_C had a better fit.



(a) (b)

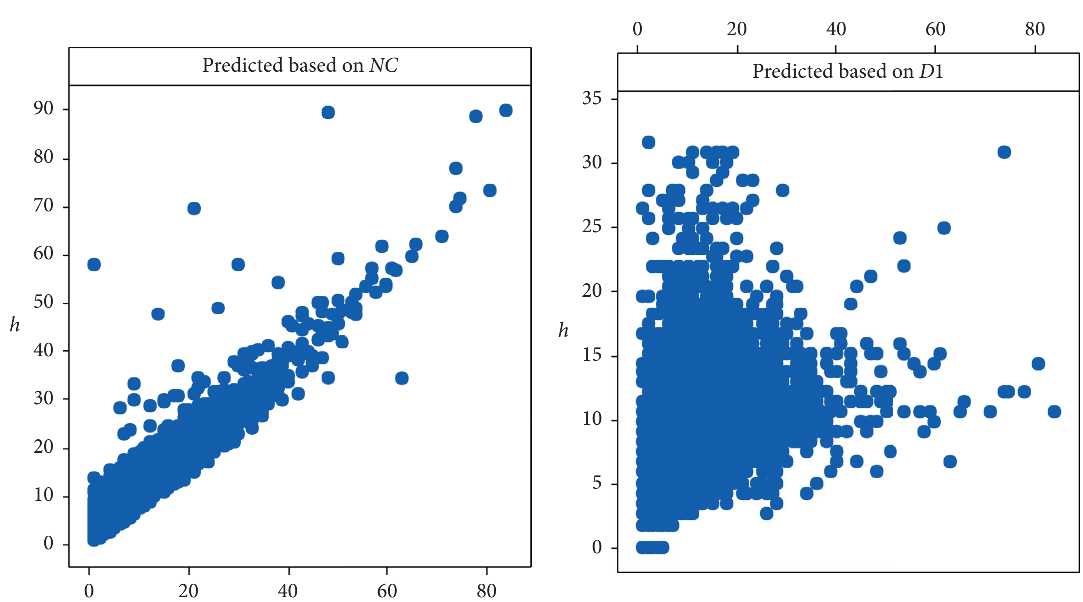
Figure 1: Plot of the fitted curve with data SNLR models.

Table 4: The parameter estimates of SLR models for N_C and D_I.

		Unstandardized coefficients		Standardized coefficients	T	p	95.0% confidence interval for B	
		B	Std. error	Beta			Lower bound	Upper bound
1	(Constant)	4.429	0.023		191.668	<0.001	4.384	4.475
	N_C	0.005	<0.001	0.754	197.221	<0.001	0.005	0.005
2	(Constant)	0.172	0.047		3.629	<0.001	0.079	0.265
	D_I	0.797	0.006	0.633	140.391	<0.001	0.786	0.809

Table 5: Five-fold cross-validation to compare SNLR and SLR models.

Method	Factor	R^2	RMSE	MAE
SLR	N_C	0.928	1.523	1.285
	D_I	0.402	4.532	3.897
SNLR	N_C	0.547	3.342	2.843
	D_I	0.243	6.517	5.761



(a) (b)

Figure 2: Plot of actual (h) values versus predicted values for SNLR models based on N_C and D_I.

Table 6: Pearson and Spearman correlations between actual values and predicted values.

	Spearman's rho		Pearson	
	Correlation coefficient	p	Correlation coefficient	p
Predicted values (based on N_C)	0.954	<0.001	0.967	<0.001
Predicted values (based on D_I)	0.779	<0.001	0.632	<0.001

Table 7: The results of the MNLR model to study the effect of N_C and D_I on h.

Factor	Source	Sum of squares	df	Mean squares	F	R^2	p
N_C, D_I	Regression	1895013.156	7	270716.1652	126717.88	0.936	<0.001
	Residual	62943.8438	29463	2.136369134			
	Uncorrected total	1957957	29470				
	Corrected total	987069.2904	29469				

Table 8: The parameter estimates of the MNLR model.

Parameter	Estimate	Std. error	95% confidence interval		p
			Lower bound	Upper bound	
b_1	0.673	0.020	0.633	0.712	<0.001
b_2	0.419	0.014	0.392	0.445	<0.001
b_3	−0.183	0.028	−0.238	−0.128	<0.001
b_4	0.939	0.061	0.819	1.058	<0.001
b_5	0.129	0.028	0.073	0.184	<0.001
b_6	0.424	0.027	0.372	0.477	<0.001
b_7	0.370	0.084	0.206	0.534	<0.001

Table 9: The parameter estimates of the MLR models for N_C and D_I.

Model		Unstandardized coefficients		Standardized coefficients	t	p	95.0% confidence interval for B	
		B	Std. error	Beta			Lower bound	Upper bound
1	(Constant)	1.003	0.033		30.751	<0.001	0.939	1.067
	N_C	0.004	<0.001	0.606	184.821	<0.001	0.004	0.004
	D_I	0.528	0.004	0.419	127.680	<0.001	0.520	0.536
	N_C* D_I	-0.132	0.014	-0.311	-9.429	<0.001	-0.182	-0.082

3.2. MNLR Results. This part is to study the simultaneous impacts of N_C and D_I on h. Tables 7 and 8 summarize the results of the MNLR model. As Table 7 indicates, the N_C and D_I factors had a significant effect on h ($p < 0.001$). Table 8 shows the parameter estimates of the MNLR model.

Based on the results of Table 8, we can estimate h as a function of NC and D1, by

$$\widehat{h}_{N_C,D_1} = 0.673N_C^{0.419} - 0.183D_1^{0.939} + 0.129N_C^{0.424}D_1^{0.370}. \tag{9}$$

Also, the R^2 value for this equation to estimate h was 93.6% that is not significantly greater than 93.4% ($\widehat{h}_C$).

Table 9 shows the results of MLR as a comparative method. Based on the results of Table 9, we can estimate h as a function of N_C and D_I, by

$$\widehat{h}_{N_C,D_1} = 1.003 + 0.004N_C + 0.528D_1 - 0.0.132N_CD_1, \tag{10}$$

respectively. Also, the R^2 value for this equation to estimate h was 72.2%. As it can be observed, the MNLR method acts better than the MLR method.

Table 10 summarizes the results of five-fold cross-validation. The results confirm that the SNLR method acts better than the SLR method.

Figure 3 and Table 11 show the plot of actual values versus predicted values and the correlations between them.

Table 10: five-fold cross-validation to compare MNLR and MLR models.

Method	Factor	R^2	RMSE	MAE
MLR	N_C	0.942	1.387	1.016
MNLR	N_C	0.717	2.419	2.013

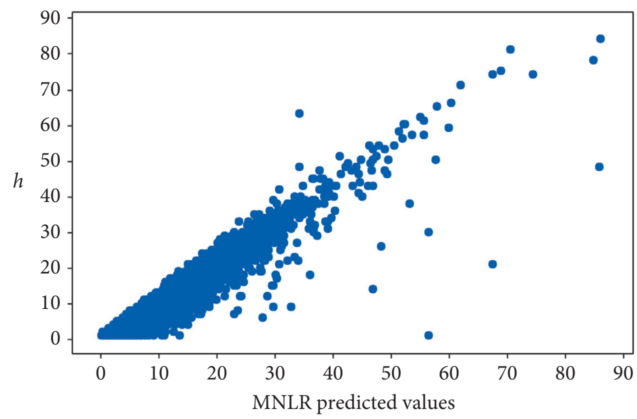


Figure 3: Plot of actual values versus predicted values for MNLR model.

Table 11: Pearson and Spearman correlations between actual values and predicted values for the MNLR model.

	Spearman's rho		Pearson	
	Correlation coefficient	p	Correlation coefficient	p
Predicted values (based on N_C and D_1)	0.968	<0.001	0.953	<0.001

Table 12: The parameter estimates of the MNLQR model.

Category	Parameter	Estimate	Std. error	95% confidence interval		p
				Lower bound	Upper bound	
1	b_1	0.929	0.010	0.909	0.948	<0.001
	b_2	0.230	0.008	0.214	0.246	<0.001
	b_3	0.104	0.020	0.064	0.144	<0.001
	b_4	0.813	0.079	0.658	0.968	<0.001
	b_5	−0.057	0.015	−0.087	−0.027	<0.001
	b_6	0.322	0.020	0.284	0.361	<0.001
	b_7	0.729	0.079	0.574	0.883	<0.001
2	b_1	11.837	322.351	−620.073	643.748	<0.001
	b_2	0.211	1.558	−2.844	3.266	<0.001
	b_3	−4.951	14.521	−33.416	23.514	<0.001
	b_4	0.021	0.182	−0.335	0.377	<0.001
	b_5	−5.983	335.988	−664.626	652.661	<0.001
	b_6	0.288	1.739	−3.122	3.698	<0.001
	b_7	−0.006	0.256	−0.508	0.496	<0.001
3	b_1	1.682	0.191	1.307	2.057	<0.001
	b_2	0.319	0.036	0.247	0.390	<0.001
	b_3	0.082	0.194	−0.297	0.462	<0.001
	b_4	0.554	0.564	−0.552	1.659	<0.001
	b_5	−0.069	0.051	−0.168	0.030	<0.001
	b_6	0.672	0.057	0.561	0.783	<0.001
	b_7	0.093	0.036	0.022	0.164	<0.001
4	b_1	0.414	0.032	0.352	0.476	<0.001
	b_2	0.523	0.009	0.505	0.541	<0.001
	b_3	4.709	0.643	3.449	5.969	<0.001
	b_4	−0.494	0.101	−0.693	−0.295	<0.001
	b_5	−0.001	0.001	−0.003	0.000	<0.001
	b_6	1.275	0.055	1.167	1.383	<0.001
	b_7	−1.348	0.148	−1.637	−1.059	<0.001

As can be seen, the MNLR model can nicely estimate the values of h.

3.3. MNLQR Results. This part is to study the simultaneous impacts of N_C and D_1 on different quartiles of h. Based on the results of Table 12, we can conclude that the N_C and D_1 factors had a significant effect on h ($p < 0.001$), in every category. Based on the results, h can be estimated as a function of N_C and D_1, by

$$\widehat{h}_{N_C,D_1} = b_1 N_C^{b_2} + b_3 D_1^{b_4} + b_5 N_C^{b_6} D_1^{b_7}, \qquad (11)$$

in categories 1 to 4, respectively.

4. Conclusion

This research dealt with a statistical approach to model the h -index (h) based on the total number of citations (N_C) and the duration from the publishing of the first article (D_1). To determine the effect of every factor (N_C and D_1) on h, we applied a set of simple nonlinear regression. The results indicated that both N_C and D_1 had a significant effect on h ($p < 0.001$) and we can estimate h as a function of N_C and D_1, by

$$\widehat{h}_{N_C} = 0.600 N_C^{0.476}, \qquad (12)$$

$$\widehat{h}_{D_1} = 0.667 D_1^{1.041}, \qquad (13)$$

respectively. Also, the R^2 values of these equations to estimate h was 93.4% and 39.8%, respectively, which verified that the model based on N_C had a better fit.

Then, to record the simultaneous effects of N_C and D_1on h, multiple nonlinear regression was applied. The results indicated that N_C and D_1 had a significant effect on h ($p < 0.001$) and we can estimate h as a function of N_C and D_1, by

$$\widehat{h}_{N_C,D_1} = 0.673 N_C^{0.419} - 0.183 D_1^{0.939} + 0.129 N_C^{0.424} D_1^{0.370}. \qquad (14)$$

Also, the R^2 value of this equation to estimate h was 93.6% that was not significantly greater than 93.4% ($\widehat{h}_C$).

Finally, to model and estimate h, as a function of N_C and D_1, the multiple nonlinear quartile regression was used. The goodness of the fitted model was also assessed.

As an important result, since the h-index is significantly affected by D_1, it is suggested to adjust the h-index based on D_1 or the times that the papers are published. Moreover, because the previous studies have verified the impact of the number of authors (N_A) of the papers on the h-index, hence it is also suggested to adjust the h-index based on N_A. As a good path for future works, the authors suggest defining a new measure as

$$\sum_{i=1}^{n} \frac{1}{(D_i + 1)} \frac{N_{C,i}}{N_{A,i}}, \qquad (15)$$

to measure the productivity of researchers, where n is the number of papers, D_i is the time when the paper i has been published (based on years), $N_{C,i}$ is the number of citations for the paper i, and $N_{A,i}$ is the number of authors for the paper i.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] J. E. Hirsch, "An index to quantify an individual's scientific research output," in *Proceedings of the National Academy of Sciences*, vol. 102, no. 46, pp. 16569–16572, 2005.

[2] T. Braun, W. Glänzel, and A. Schubert, "A Hirsch-type index for journals," *Scientometrics*, vol. 69, no. 1, pp. 169–173, 2006.

[3] A.-W. Harzing and R. van der Wal, "A Google Scholar h-index for journals: an alternative metric to measure journal impact in economics and business," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 1, pp. 41–46, 2009.

[4] A. Schubert and W. Glänzel, "A systematic analysis of Hirsch-type indices for journals," *Journal of Informetrics*, vol. 1, no. 3, pp. 179–184, 2007.

[5] P. Vinkler, "Eminence of scientists in the light of theh-index and other scientometric indicators," *Journal of Information Science*, vol. 33, no. 4, pp. 481–491, 2007.

[6] S. Ruch and R. Ball, "Various correlations between the H-Index and citation rate (CPP) in neuroscience and quantum physics: new findings," *International Journal of Information Science and Management*, vol. 8, no. 1, pp. 1–19, 2010.

[7] A. Yong, "A critique of hirsch's citation index: a combinatorial fermi problem," *Notices of the American Mathematical Society*, vol. 61, no. 9, pp. 1040–1050, 2014.

[8] J. Bar-Ilan, "Rankings of information and library science journals by JIF and by h-type indices," *Journal of Informetrics*, vol. 4, no. 2, pp. 141–147, 2010.

[9] L. Egghe and R. Rousseau, "An informetric model for the Hirsch-index," *Scientometrics*, vol. 69, no. 1, pp. 121–129, 2006.

[10] Q. L. Burrell, "Formulae for the h-index: a lack of robustness in Lotkaian informetrics?," *Journal of the American Society for Information Science and Technology*, vol. 64, no. 7, pp. 1504–1514, 2013.

[11] W. Glänzel, "On the h-index-a mathematical approach to a new measure of publication activity and citation impact," *Scientometrics*, vol. 67, no. 2, pp. 315–321, 2006.

[12] J. E. Iglesias and C. Pecharromán, "Scaling the h-index for different scientific ISI fields," *Scientometrics*, vol. 73, no. 3, pp. 303–320, 2007.
[13] L. Egghe, L. Liang, and R. Rousseau, "A relation between h-index and impact factor in the power-law model," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 11, pp. 2362–2365, 2009.
[14] A. Bletsas and J. N. Sahalos, "Hirsch index rankings require scaling and higher moment," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 12, pp. 2577–2586, 2009.
[15] L. Egghe and R. Rousseau, "The Hirsch index of a shifted Lotka function and its relation with the impact factor," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 5, pp. 1048–1053, 2012.
[16] L. Bertoli-Barsotti and T. Lando, "On a formula for the h-index," *Journal of Informetrics*, vol. 9, no. 4, pp. 762–776, 2015.
[17] J. E. Hirsch, "Does the h index have predictive power?," *Proceedings of the National Academy of Sciences*, vol. 104, no. 49, p. 19193, 2007.
[18] H. Haghbin, M. R. Mahmoudi, and Z. Shishebor, "Large sample inference on the ratio of two independent binomial proportions," *Journal of Mathematical Extension*, vol. 5, no. 1, pp. 87–95, 2011.
[19] M. R. Mahmoudi, J. Behboodian, and M. Maleki, "Large sample inference about the ratio of means in two independent populations," *Journal of Statistical Theory and Applications*, vol. 16, no. 3, pp. 366–374, 2017.
[20] M. R. Mahmoudi and M. Mahmoodi, "Inferrence on the ratio of variances of two independent populations," *Journal of Mathematical Extension*, vol. 7, no. 2, pp. 83–91, 2014.
[21] M. R. Mahmoudi and M. Mahmoodi, "Inferrence on the ratio of correlations of two independent populations," *Journal of Mathematical Extension*, vol. 7, no. 4, pp. 71–82, 2014.
[22] M. R. Mahmoudi, R. Nasirzadeh, and M. Mohammadi, "On the ratio of two independent skewnesses," *Communications in Statistics-Theory and Methods*, vol. 48, no. 7, pp. 1721–1727, 2019.
[23] M. R. Mahmoudi, M. Mahmoudi, and E. Nahavandi, "Testing the difference between two independent regression models," *Communications in Statistics-Theory and Methods*, vol. 45, no. 21, pp. 6284–6289, 2016.
[24] M. R. Mahmoudi, M. Maleki, and A. Pak, "Testing the equality of two independent regression models," *Communications in Statistics-Theory and Methods*, vol. 47, no. 12, pp. 2919–2926, 2018.
[25] M. R. Mahmoudi, "On comparing two dependent linear and nonlinear regression models," *Journal of Testing and Evaluation*, vol. 47, no. 1, pp. 449–458, 2018.
[26] P. Ji-jun, M. R. Mahmoudi, D. Baleanu, and M. Maleki, "On comparing and classifying several independent linear and non-linear regression models with symmetric errors," *Symmetry*, vol. 11, no. 6, p. 820, 2019.
[27] M. Bahrami, M. J. Amiri, M. R. Mahmoudi, and S. Koochaki, "Modeling caffeine adsorption by multi-walled carbon nanotubes using multiple polynomial regression with interaction effects," *Journal of Water and Health*, vol. 15, no. 4, pp. 526–535, 2017.
[28] M. R. Mahmoudi, M. H. Heydari, and K.-H. Pho, "Fuzzy clustering to classify several regression models with fractional Brownian motion errors," *Alexandria Engineering Journal*, vol. 59, no. 4, pp. 2811–2818, 2020.
[29] M. R. Mahmoudi, M. Mahmoudi, and A. Pak, "On comparing, classifying and clustering several dependent regression models," *Journal of Statistical Computation and Simulation*, vol. 89, no. 12, pp. 2280–2292, 2019.
[30] M. R. Mahmoudi, M. Maleki, and A. Pak, "Testing the difference between two independent time series models," *Iranian Journal of Science and Technology, Transactions A: Science*, vol. 41, no. 3, pp. 665–669, 2017.
[31] M. R. Mahmoudi, M. H. Heydari, and R. Roohi, "A new method to compare the spectral densities of two independent periodically correlated time series," *Mathematics and Computers in Simulation*, vol. 160, pp. 103–110, 2019.
[32] M. R. Mahmoudi, M. H. Heydari, and Z. Avazzadeh, "Testing the difference between spectral densities of two independent periodically correlated (cyclostationary) time series models," *Communications in Statistics-Theory and Methods*, vol. 48, no. 9, pp. 2320–2328, 2019.
[33] M. H. Heydari, Z. Avazzadeh, and M. R. Mahmoudi, "Chebyshev cardinal wavelets for nonlinear stochastic differential equations driven with variable-order fractional Brownian motion," *Chaos, Solitons & Fractals*, vol. 124, pp. 105–124, 2019.
[34] M. R. Mahmoudi and M. Maleki, "A new method to detect periodically correlated structure," *Computational Statistics*, vol. 32, no. 4, pp. 1569–1581, 2017.
[35] A. R. Nematollahi, A. R. Soltani, and M. R. Mahmoudi, "Periodically correlated modeling by means of the periodograms asymptotic distributions," *Statistical Papers*, vol. 58, no. 4, pp. 1267–1278, 2017.
[36] M. R. Mahmoudi, M. H. Heydari, and Z. Avazzadeh, "On the asymptotic distribution for the periodograms of almost periodically correlated (cyclostationary) processes," *Digital Signal Processing*, vol. 81, pp. 186–197, 2018.
[37] M. R. Mahmoudi, M. H. Heydari, Z. Avazzadeh, and K.-H. Pho, "Goodness of fit test for almost cyclostationary processes," *Digital Signal Processing*, vol. 96, p. 102597, 2020.
[38] M. R. Mahmoudi, M. Maleki, K. Borodin, K.-H. Pho, and D. Baleanu, "On comparing and clustering the spectral densities of several almost cyclostationary processes," *Alexandria Engineering Journal*, vol. 59, no. 4, pp. 2555–2565, 2020.
[39] R. Zhou, M. R. Mahmoudi, S. N. Q. Mohammed, and K.-H. Pho, "Testing the equality of the spectral densities of several uncorrelated almost cyclostationary processes," *Alexandria Engineering Journal*, vol. 59, no. 5, pp. 3545–3550, 2020.
[40] M. R. Mahmoudi, D. Baleanu, B. Anh Tuan, and K.-H. Pho, "A novel method to detect almost cyclostationary structure," *Alexandria Engineering Journal*, vol. 59, no. 4, pp. 2339–2346, 2020.
[41] R. Roohi, M. H. Heydari, M. Aslami, and M. R. Mahmoudi, "A comprehensive numerical study of space-time fractional bioheat equation using fractional-order Legendre functions," *The European Physical Journal Plus*, vol. 133, p. 412, 2018.
[42] M. R. Mahmoudi, A. R. Nematollahi, and A. R. Soltani, "On the detection and estimation of the simple harmonizable processes," *Iranian Journal of Science and Technology (Sciences)*, vol. 39, no. 2, pp. 239–242, 2015.
[43] M. Maleki, M. R. Mahmoudi, D. Wraith, and K. H. Pho, "Time series modelling to forecast the confirmed and recovered cases of COVID-19," *Travel Medicine and Infectious Disease*, vol. 37, p. 101742, 2020.
[44] A. R. Abbasi, M. R. Mahmoudi, and Z. Avazzadeh, "Diagnosis and clustering of power transformer winding fault types by cross-correlation and clustering analysis of FRA results," *IET Generation, Transmission & Distribution*, vol. 12, no. 19, pp. 4301–4309, 2018.