Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Evaluation of the differences in competition strength of high-level table tennis matches: a case study based on a multiple regression model

Xia Chen[1], Huagen Yin[2,✉], Yanxiang Zhou[3], Lin Zhou[4]

[1] *School of Physical Education, Putian University, Putian, Fujian, 351100, China*

[2] *College of Physical Education, Shangrao Normal University, Shangrao, Jiangxi, 334001, China*

[3] *Shangrao Health Vocational College, Shangrao, Jiangxi, 334600, China*

[4] *East University of Heilongjiang, Harbin, Heilongjiang, 610043, China*

ABSTRACT

With the reforms in competition rules and equipment by the International Table Tennis Federation (ITTF), the number of rounds in table tennis matches has increased, placing higher demands on athletes' abilities to transition between technical and tactical offensive and defensive strategies, as well as on their physical and psychological qualities. Therefore, this study employs methods such as the strength difference evaluation, competition performance (CP), and multiple regression. Using 48 international important matches in which Player W (anonymous) participated as case studies, the study evaluates and predicts the competitive performance of W when facing athletes of different world rankings. The results indicate that in matches against athletes with different skill rankings, the phases where the technical strength difference significantly affects competition performance (CP) are the attack-after-serve phase and attack-after-receive phase, followed by the rally phase, and finally, the rally phase Ⅱ. The competitive level in serve rounds is superior to that in receive rounds. The Kruskal-Wallis test results reveal significant fluctuations in the competition performance (CP) during the rally phase Ⅱ, demonstrating highly significant differences ($P < 0.01$). In matches, Player W has a very low probability of winning when not holding an absolute advantage in key techniques (the first four strokes) - particularly pronounced when facing athletes ranked in the world top 20. The multiple regression model for the technical strength difference in table tennis matches plays a certain role in predicting the performance of athletes in terms of technical and tactical indicators during matches, offering a clear reflection of the effectiveness of these indicators.

Keywords: table tennis, strength difference, competition performance, multiple regression, Four-phase-index in two rounds

✉ Corresponding author.
E-mail address: yinhuagen1990@163.com (H. Yin).

Received 25 October 2024; Revised 10 December 2024; Accepted 05 March 2025; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-088

1. Introduction

The athletic level of Chinese table tennis players has long been at the top level of the world, and the Chinese table tennis team has performed particularly brightly on the world stage, and its achievements have taken the lead, with the emergence of many Grand Slam players one after another. The reason why the Chinese table tennis team has been able to maintain a long-lasting position in the world is, on the one hand, the fact that Chinese table tennis players have advanced technology, a technical system superior to that of every country, and the spirit of the Long March, which is characterized by perseverance and the courage to suffer [1]. On the other hand, it is the long time efforts of the scientific research team behind the Chinese table tennis team. While the athletes are fighting on the field, the scientific research team is working hard behind the scene to evaluate the strength of the opponents [2]. The main work of the scientific research team is to collect and study the technical and tactical data of Chinese athletes and athletes of foreign associations, adjust the technical and tactical characteristics of their own athletes in response to the technical and tactical changes of their opponents, and absorb the technical and tactical advantages of outstanding athletes of foreign associations, and then incorporate them into their own athletes, thus optimizing the techniques to make up for their own shortcomings [3-6]. Technological innovation is the most direct way to improve athletes' competitive strength. Although the Chinese table tennis team is the top level in the world, it has the courage to admit its own shortcomings and dare to innovate and make trial and error, and this spirit is one of the factors that make the Chinese table tennis team everlasting [7-9]. In addition, after the reform of the table tennis system led by the International Federation of Armed Forces (IFAB), many famous players have emerged in the world of table tennis, and each athlete has different technical and tactical characteristics and playing styles, and has shown excellent technical strength, which has caused a great impact on the daily training mode of Chinese table tennis team [10-11]. Analyzing the playing styles of different athletes and iteratively optimizing them in terms of technique and tactics are more valuable for the cultivation of Chinese table tennis reserve talents [12].

A large number of scholars have analyzed the changes in ball speed, power, rotation, and landing point embodied by athletes in table tennis matches, assessing the strength differences between athletes from the level of tactics and techniques chosen by the athletes, and assisting in the further improvement of athletes' competitive level. Literature [13] shows that the large-scale high-dimensional data in table tennis matches poses a high challenge to tactical analysis, and proposes a visual analysis system, Tac-Miner, to provide support for the visual analysis of multiple matches through dimensionality reduction algorithms as well as interactive symbolic analysis. Literature [14] constructed a table tennis tactical analysis system based on data mining technology, which can carry out data query and statistics in a targeted manner and form a systematic analysis of table tennis tactics using information processing technology, and the proposed system has an outstanding contribution to the improvement of table tennis player's training level and competitive quality. Literature [15] introduced the basic posture strategy analysis model to calculate the usage rate and scoring rate of common tactics used by athletes in table tennis matches, which provides reliable data support for analyzing the tactical behaviors and strength differences of athletes with different racket hands. Literature [16] proposed an interactive table tennis visualization system to visualize and analyze table tennis matches from three perspectives: time-oriented analysis, statistical analysis and tactical analysis, so as to gain insights into the athletes' technical and tactical levels in terms of hitting movements, hitting positions and other fine-grained aspects. Literature [17] established a three-layer BP neural network model for the technical and tactical evaluation of table tennis matches, which utilizes the model's nonlinear dynamic processing capability and high fitting accuracy to calculate athletes' technical and tactical indexes and the probability of winning matches, and can also be based on which to provide targeted training programs for players. Literature [18] uses target tracking and trajectory estimation methods to establish a table tennis video game tactical analysis platform, which

recognizes table tennis balls in the game video through image fuzzy edge recognition algorithms, and records its speed, landing point and other key data information to help athletes develop more effective tactics and strategies.

This paper makes a predictive statistical analysis on the construction and evaluation model of the competitive strength of W and the athletes of different rankings in terms of strength in the world (top 10, top 11 to 20, and 20 or above). This statistical method is the first detailed analysis of both technique and tactics levels of playing table tennis in the field of diagnosis and analysis of sports techniques, aiming at enriching and perfecting the statistical methods of table tennis techniques and tactics, and providing a theoretical reference for furthering the scientific development of the technique level of table tennis.

2. Research Object and Methods

2.1. Selection of Match Videos

This paper takes W, the new-generation key athlete of table tennis, as the research object. The player employs a right-handed shake-hands grip and exhibits a playing style characterized by two-winged looping and fast-break techniques. His playing method is based on the combination of two arcs near the table and the placement is tricky. He is the youngest champion of the Tour Finals in the history of the International Table Tennis Federation and has defeated top Chinese table tennis athletes in important international events many times. In this regard, the 48 important matches that W participated in from 2018 to 2021 were selected for research. Among them, a total of 19, 17, and 12 matches were against the top 10, top 11-20, and 20 or above athletes in the world, respectively. Before the study, the data of the respondents were collected from various match videos published on official websites such as the International Table Tennis Federation, CCTV and Bilibili. Since all the video of the match is from the television broadcast on the official website and the Internet, the study does not need to be approved by the ethics committee of the local institution, according to the regulations of the national legislature.

2.2. Methods

2.2.1. Video observation. All videos in this study are from the match videos authorized by the official website of Table tennis for public play, and they are all selected from the World Championships, World Cup, ITTF World Tour, Japan National Table Tennis Championship and other representative important events. The scoring of both athletes in the matches was based on the "Four-phase-index Evaluation Method" constructed, and the score and loss of one athlete's last stroke were the observation points. The Excel-based Four-phase-index Evaluation Table was used to make statistics on the original data of the athletes on both sides. For example, in the match W vs. Ma Long, W and his serve score in the 7th stroke should be included in that in rally phase I; Ma Long's receive should be included in the score of the 8th stroke in rally phase II, so it shall be included in the score of the 8th stroke in rally phase II, and so on. The specific strokes included in the Four-phase-index of two rounds are the serve round (attack-after-serve <serve, the 3rd stroke, the 5th stroke lost>, rally phase I <the 5th stroke scored, the 7th stroke and the subsequent odd strokes>) and the receive round (attack-after-receive <receive, the 4th stroke>, rally phase II <the 6th stroke and the subsequent even strokes>), and the table tennis "Four-phase-index in two rounds " technical and tactical system model is shown in Figure 1. Moreover, we collected all the data independently to ensure the authenticity and reliability of the data. We further trained three students majoring in table tennis to make auxiliary observation records. Compared with the data collected by the three students, the Kappa value is 0.921 through Kappa Test, indicating that the observation data are consistent well.

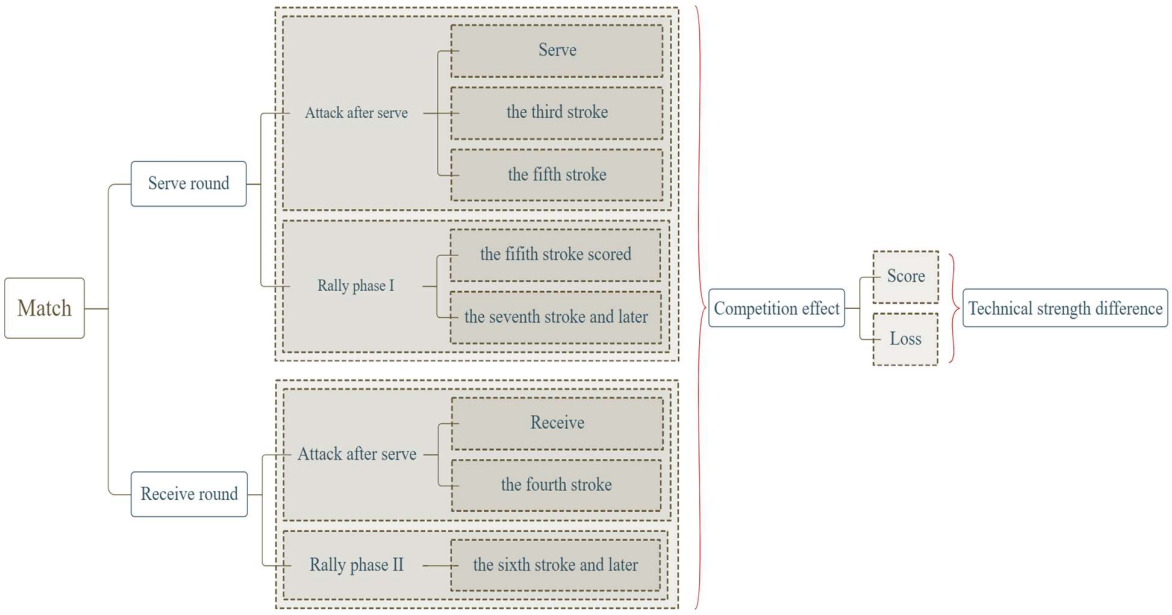

Fig. 1. Four-phase-index in two rounds Technical and tactical system model

2.2.2. Strength difference evaluation. The essence of strength difference (SD) is the difference between the percentage of one athlete's score and opponent's in the phase (a round or a match) in the total score and loss of the whole match. SD is called Strength Difference Evaluation Method as it is used to evaluate the strength difference between the two athletes in the process of on-the-spot confrontation. This study mainly aims to evaluate the strength difference in the four phases of the research objects. The evaluation criteria are shown in Table 1. The four phases strength difference was mainly calculated based on the "Four-phase-index Evaluation Method" constructed concerning the score and loss of both athletes one by one. For example, ① attack-after-serve phase of Athlete A corresponds to the attack-after-receive phase of Athlete B; ② the attack-after-receive phase of Athlete A corresponds to the attack-after-serve phase of Athlete B; ③ the rally phase I of Athlete A corresponds to the rally phase II of Athlete B; ④ the rally phase II of Athlete A corresponds to the rally phase I of Athlete B. (Note: Athlete A refers to W, and Athlete B refers to his opponent). Additionally, for the convenience of understanding the formula for calculating the strength difference in the match, this study has established the code for scoring and losing points in table tennis matches between athletes from both sides (Table 2). Based on the observation code table established in Table 2, the match data for both sides can be converted into data for one side. Letting I represent the total points scored and lost in the match, the formula for the strength difference in each phase for athlete A can be expressed as:

$$Athlete\ A's\ Strength\ Difference\ in\ attack\text{-}after\text{-}serve\ phase(SD_{1\text{-}3}) = \frac{(A^{+}+B^{+})-(A^{-}+B^{-}+C^{-})}{I} \times 100\% \quad (1)$$

$$Athlete\ A's\ Strength\ Difference\ in\ attack\text{-}after\text{-}receive\ phase(SD_{2\text{-}4}) = \frac{(X^{+}+Y^{+})-(X^{-}+Y^{-})}{I} \times 100\% \quad (2)$$

$$Athlete\ A's\ Strength\ Difference\ in\ Rally\ Phase\ I(SD_{\geq 5}) = \frac{(C^{+}+D^{+})-D^{-}}{I} \times 100\% \quad (3)$$

$$Athlete\ A's\ Strength\ Difference\ in\ Rally\ Phase\ II(SD_{\geq 6}) = \frac{Z^{+}-Z^{-}}{I} \times 100\% \quad (4)$$

Table 1. Strength difference evaluation of men's singles of table tennis

Indexes	Wide positive disparity	Large positive disparity	Equal strength	Large negative disparity	Wide negative disparity
Attack-after-serve phase	⩾0.070	(0.070,0.007]	(0.007,-0.050]	(-0.050,-0.100]	<-0.100
rally phase I	⩾0.100	(0.100,0.040]	(0.040,0.010]	(0.010,-0.030]	<-0.030
Attack-after-receive phase	⩾0.070	(0.070,0.027]	(0.027,-0.020]	(-0.020,-0.060]	<-0.060
rally phase II	⩾0.010	(0.010,-0.020]	(-0.020,-0.067]	(-0.067,-0.100]	<-0.100

Table 2. Table Tennis Match Tactical Scoring and Losing Points Codes

Rounds	Hits	Scoring observation points and code	Losing point and code	Total code
Serve round	serve	Opponent's fault in receiving the serve(A+)	Own serve fault (A-)	A
	3rd stroke	Opponent's fault in the 4th stroke (B+)	Own fault in the 3rd stroke (B-)	B
	5th stroke	Opponent's fault in the 6th stroke(C+)	Own fault in the 5th stroke (C-)	C
	Rally I	Opponent's fault in the 8th stroke and the subsequent strokes(D+)	Own fault in the 7th stroke and the subsequent strokes (D-)	D
Receive round	receive	Opponent's fault in the 3rd stroke(X+)	Own fault in the receive (X-)	X
	4th stroke	Opponent's fault in the 5th stroke(Y+)	Own fault in the 4th stroke (Y-)	Y
	Rally II	Opponent's fault in the 7th stroke and the subsequent strokes(Z+)	Own fault in the 6th stroke and the subsequent strokes (Z-)	Z

2.2.3. Table tennis competition performance. The score is an important performance of the athletes to show the competition performance. In the 11-score table tennis match, the maximum difference between the victory and defeat of the athletes in a match is 22 points, i.e., 11:0. At 11:0, the competition performance (CP) of the winner is 22 and that of the losing athlete is 0. The specific calculation equation is as follows:

$$CP = 22 - \sqrt{\frac{\sum_{i=1}^{n}(x_i - 11)^2}{n}} \tag{5}$$

In equation (5), x_i is the difference of the scores in each match, n is the inning of matches. In addition, equation (5) was calculated according to the theoretical winning scores of each match and the actual net winning scores of each match. For this reason, some special circumstances that an athlete wins the whole match, but loses to the opponent on the total score (for example, when the scores of a match are 11:7, 11:8, 3:11, 5:11, 11:9, and 11:9, respectively, the winner wins with a big margin of 4:2, but the individual total score is 52/55 less than the opponent) will also cause the CP value of the winner to be less than that of the loser. According to relevant research, the probability of this situation in international matches is about 5%.

2.2.4. Mathematical statistics. A statistical software (Microsoft Excel) was used to establish a Four-phase-index Evaluation Table to count the scores and losses of various technical indexes. SPSS 25.0 (SPSS Inc., Chicago, IL, USA) was employed to carry out multiple linear regression analysis on the processed data while taking the Competition Performance (CP) as the dependent variable, and the strength difference of the four indicators (attack-after-serve phase, rally phase I, attack-after-receive phase, and rally phase II) as the independent variables. Then, we analyzed the contribution of independent variables to the dependent variable. Finally, the nonparametric Kruskal Wallis Test of multiple independent samples was used to test the Four-phase-index strength difference between W and the table tennis athletes ranking the top 10, top 11-20, and beyond 20 in the world.

3. Results and Analysis

3.1. SD Evaluation between W and the Top 10 World Athletes

Table 3 shows the statistical results of W's competition strength difference in matches with the top 10 world athletes. On the whole, the competition strength of W's attack-after-serve phase, rally phase I, and attack after receive phase, fall within the scope of "equal strength" ($\overline{X}$ = 0.001, $\overline{X}$ = 0.027, $\overline{X}$ = 0.010), and the competition strength of rally phase II pertains to "large negative disparity" ($\overline{X}$ = -0.086). The results show that in the match with the top 10 world athletes, W's competitive level in rally phase II (after the 6th stroke) is relatively poor, which is inferior to that in attack-after-serve phase, rally phase I, and attack-after-receive phase.

It can also be seen from Table 3 that only one index of W is evaluated as Wide Negative Disparity (No. 2, 3, 4, 6, 13 and 18) in the match, indicating that he lost all these matches. For other matches that he lost (No. 5, 10, 12, 15, 17 and 19), at least two indicators are evaluated as or lower than "equal strength". For example, in the match between W and Chinese athletes X (No. 5) and M (No. 10), even if one index reached a "large positive disparity", the other three indexes reached "equal strength", and the four indexes were "equal strength", he still lost the match. It is obvious that W is difficult to win the match among the top 10 world athletes without the absolute advantage of the competitive strength of a certain index.

Judging from the matches won by W (No. 1, 7, 8, 9, 11, 14 and 16), except for those against Chinese athlete M (No. 9), at least one index of the other winning matches has reached the level of "wide positive disparity". In the match with Chinese athlete M (No. 9), although W won the match with two indicators reaching a "large positive disparity" (4:2), he still lagged behind his opponents (49/52) with a small margin, which is related to the fact that W did not have absolute advantage of the competitive strength of a certain index. What's more, it is worth noting that W's dominant techniques are mainly concentrated in the attack after serve phase, and the attack after receive phase. On the other hand, his techniques were not stably wielded in the rally phases I and II, which may be related to his playing style (valuing the quality of attack after serve and attack after receive). From the characteristics of today's large ball matches, there will be more and more rounds in table tennis, and the techniques of rally phases have become particularly important techniques and tactics at present. This means that W's technique and tactic system needs to be improved.

Table 3. list of strength difference evaluation of W in the match with the top 10 world athletes

S/N	Athletes	Attack-after-serve phase		Rally phase I		Attack-after-receive phase		Rally phase II		Results
		SD	Evaluation	SD	Evaluation	SD	Evaluation	SD	Evaluation	
1	F	0.130	++	0.013	0	0.000	0	-0.078	-	3:1(41/36)
2	F	0.011	+	0.000	0	-	-	-	--	1:4(38/50)

						0.034		0.114		
3	F	0.049	+	-0.029	-	0.010	0	-0.186	--	2:4(43/59)
4	X	-0.110	--	0.028	0	-0.028	-	-0.028	0	2:4(47/62)
5	X	-0.047	0	0.016	0	0.031	+	-0.039	0	2:4(61/66)
6	X	0.007	0	0.028	0	0.056	+	-0.112	--	3:4(70/73)
7	Z	0.203	++	0.087	+	0.043	+	-0.058	0	4:0(44/25)
8	Z	-0.087	-	0.007	-	0.145	++	-0.007	+	4:3(73/65)
9	M	0.010	+	0.069	+	-0.049	-	-0.059	0	4:2(49/52)
10	M	-0.069	-	0.026	0	0.000	0	-0.052	0	1:4(52/63)
11	M	-0.009	0	0.104	++	0.047	+	-0.066	-	4:2(56/50)
12	M	-0.008	0	0.034	0	0.000	0	-0.050	0	3:4(58/61)
13	L	-0.057	-	0.034	0	-0.126	--	-0.046	0	1:4(35/52)
14	L	0.000	0	0.000	-	0.107	++	0.000	+	4:1(57/46)
15	L	-0.047	0	0.028	0	-0.047	-	-0.470	0	2:4(47/59)
16	J	0.079	++	0.000	-	0.114	++	-0.038	0	4:0(45/34)
17	J	0.018	+	0.000	-	0.009	0	-0.071	-	2:4(54/59)
18	B	-0.011	0	0.042	+	-0.138	--	-0.064	0	1:4(39/55)
19	K	-0.043	0	0.026	0	0.017	0	-0.094	-	3:4(53/64)
$\bar{X}$		0.001	0	0.027	0	0.010	0	-0.086	-	

Note: in the table, "0" indicates that the competition strength is evaluated as "equal strength", "+" indicates "large positive disparity", "+ +" indicates "wide positive disparity", "-" indicates "large negative disparity", "--" indicates "wide negative disparity", and the same in the following table.

3.2. Multiple Regression Analysis of W's Match with the Top 10 World Athletes

In this study, X1 (competition strength in the attack after serve phase), X2 (competition strength in the rally phase I), X3 (competition strength in the attack after receive phase) and X4 (competition strength in the rally phase II) were taken as independent variables and Y (competition performance "CP") was taken as a dependent variable. A multiple regression model was established by SPSS statistical software. W and the top 10 world athletes in 19 matches were imported into the software, and clicked "enter".

According to the results in Table 4, the complex correlation coefficient R, determination coefficient R^2, and adjustment coefficient R^2 of the multiple regression model are all above 0.82, and the values of determination coefficient and adjustment coefficient are relatively close, indicating that the data fit well with high effectiveness. The residual autocorrelation (Durbin-Watson value) of the model is 1.527, indicating that the residuals of the model are independent of each other. The analysis of variance showed that the F value of the regression model was 22.626, $P < 0.01$, indicating that the regression model is significant as a whole.

Table 4. Multiple regression model and analysis of variance results of W's match with the top 10 world athletes

R	R^2	Adjusted R^2	Standard error	Durbin-Watson	F value of regression equation	Sig value of regression equation
0.931	0.866	0.828	0.983	1.527	22.626	0.000

Table 5 shows the coefficient test results of the multiple regression model. The four independent variables, except for the significance of X4 (rally phase II) strength difference ($t = 0.906$, $Sig = 0.38$) > 0.05, showed that there is no significant difference between X4 (rally phase II) strength difference and competition performance (CP), nor is there any statistical causality, so it would be excluded from the model. The other three independent variables were significantly different from the competition performance (CP) ($P < 0.05$). Among them, there was a very significant difference between X1 and X3 and the competition performance (CP) ($P < 0.01$). It demonstrates that there is a linear relationship between X1, X2, X3 and competition performance (CP). In addition, in the collinearity diagnostic test of variables, the tolerance is between 0.846 ~ 0.939, and the value of variance inflation factor (VIF) is 1.065~1.182 < 2, indicating that the parameters in the regression model are reliable and the equation is established Xu, 2018; Huang et al., 2021). Therefore, the regression model of W's match with the top 10 world athletes is as follows:

$$Y = 9.313 + 15.026 * X1 + 20.390 * X2 + 21.747 * X3$$

The partial regression coefficient in the regression equation is the ratio that the independent variable causes the change of the dependent variables when it is not disturbed by other independent variables. The standard regression coefficient is the dimensionless quantification of the partial regression coefficient, which can be used to determine the influence of the independent variable on the dependent variable . The competition strength in the four indicators can reflect the difference in the competitive level of athletes. As a result, the standard regression coefficient of the strength difference of each phase in this study can also represent the influence of various technical and tactical abilities on the CP. The greater the standard regression coefficient of a certain strength difference, the greater the influence on the dependent variable. Hence, it can be seen from Table 5 that the CP influence of W and the top 10 world athletes in different phases is as below (in descending order): X3 (competition strength in attack-after-receive phase) > X1 (competition strength in attack-after-serve phase) > X2 (competition strength in rally phase I) > X4 (competition strength in rally phase II).

Table 5. Results of model coefficient test for W and his match with the top 10 world athletes

Model	Non standardized coefficient		Standard coefficient	t	Sig	Collinearity statistics	
	B	Standard error	Trial version			Tolerance	VIF
(constant)	9.313	0.392		23.727	0.000		
X1	15.026	3.184	0.477	4.72	0.000	0.939	1.065
X2	20.39	7.587	0.275	2.687	0.018	0.913	1.096
X3	21.747	3.48	0.665	6.25	0.000	0.846	1.182
X4	2.160	2.383	0.093	0.906	0.380	0.909	1.101

3.3. Evaluation of W's competition strength in the matches with the top 11-20 world athletes

Table 6 shows the evaluation results of W's competition strength in the matches with the top 11-20 world athletes. According to the data in the table, for the overall strength of W in 17 matches, the competition strength between the attack after serve phase and the attack after receive phase belongs to "large positive disparity" ($\overline{X}$ = 0.022, $\overline{X}$ = 0.028). The SD between rally phase I and rally phase II belongs to "equal strength" ($\overline{X}$ = 0.04, $\overline{X}$ = -0.024). These results demonstrate that W can give better play to his technical and tactical level when competing with the athletes lagging behind the world ranking, especially in the attack after serve phase and attack after receive phase.

In terms of the single match, W won 12 matches (No. 1, 2, 3, 4, 6, 8, 10, 11, 12, 14, 15 and 16), except for the weak advantage of "large positive disparity" and "equal strength" and "wide negative disparity" with Korean athlete Y (No. 8), at least one index reached the absolute advantage of "wide positive disparity" or three indexes reached the absolute advantage of "large positive disparity" in other winning matches. This shows that the possibility of winning a match without absolute advantage is not that high in terms of techniques of a phase. Putting aside the technical and tactical level, even if he wins, this is more of a narrow victory through self-regulation of psychological ability with a weak score. In W's five matches that he lost (No. 5, 7, 9, 13 and 17), except the match with German athlete O (No. 17), at least one index of the other losing matches were in the disadvantage gap of "large negative disparity" or the overall competitive strength of the four indexes is low. In addition, W in the match with German athlete O, although the strength difference in the rally phase I reached "large disparity", other indicators were at a disadvantage level, meaning that it was still difficult to win the match. This reveals that the table tennis match has a high degree of connection between techniques in different phases. It cannot simply rely on the prominence of one technique alone. It is still difficult to win if other technologies play a passive role.

Table 6. List of SD evaluation between W and the top 11-20 world athletes

S/N	Athlete	Attack after serve phase		Rally phase I		Attack after receive		Rally phase II		Results
		SD	Evaluation	SD	Evaluation	SD	Evaluation	SD	Evaluation	
1	U	0.121	++	0.060	+	-0.034	-	-0.043	0	4:2(64/52)
2	Y	0.020	+	0.061	+	0.000	0	0.010	+	4:1(54/45)
3	H	0.098	++	0.011	0	0.054	+	0.033	++	4:1(55/37)
4	U	0.100	++	0.011	0	0.056	+	-0.078	-	4:1(49/41)
5	U	-0.048	0	-0.009	-	0.009	0	-0.076	-	2:4(46/59)
6	O	0.082	++	-0.014	-	0.137	++	0.000	+	4:0(44/29)
7	G	0.034	+	-0.043	--	-0.026	-	-0.017	+	2:4(55/61)
8	Y	0.000	0	0.053	+	0.068	+	-0.106	--	4:3(67/65)
9	Y	0.009	+	0.094	+	-0.077	--	-0.051	0	3:4(57/60)
10	S	-0.039	0	0.089	+	0.049	+	0.010	++	4:2(56/45)
11	E	0.090	++	-0.022	-	0.112	++	-0.034	0	4:1(51/38)
12	E	0.067	+	0.048	+	0.038	+	-0.029	0	4:2(59/46)
13	Q	-0.009	0	-0.036	--	0.009	0	0.000	+	2:4(54/58)
14	I	-0.073	-	0.109	++	0.022	0	0.044	++	4:3(76/62)
15	R	-0.088	-	0.088	+	0.105	++	0.053	++	3:0(33/24)
16	T	0.089	++	0.069	+	0.010	0	-0.020	0	4:2(58/43)
17	O	-0.087	-	0.107	++	-0.049	-	-0.097	-	2:4(45/58)
$\bar{X}$		0.022	+	0.040	0	0.028	+	-0.024	0	

3.4. Multiple Regression Analysis of W's Match with the Top 11-20 World Athletes

The data in Table 7 display that in the regression model of W and the top 11-20 world athletes, the complex correlation coefficient R, determination coefficient R^2 and adjustment coefficient R^2 are all above 0.93, indicating that the model is of high fitting degree. Among them, the determination coefficient R^2 represents the proportion that the variation of the dependent variable can be explained by the independent variable through the regression coefficient, that is, the competition strength of the four indicators in the model can explain 95.3% change of CP. The DW value of the Durbin-Watson Test is 1.327, indicating that the residuals of the model are independent of each other. The model passed the F-test ($F = 60.639$, $P = 0.000 < 0.05$), indicating that the regression model is significant.

Table 7. Results of multiple regression model and analysis of variance for W's match with the top 11-20 world athletes

R	R^2	Adjusted R^2	Standard error	Durbin-Watson	F value of regression equation	Sig value of regression equation
0.976	0.953	0.938	0.493	1.327	61.022	0.000

According to the data results shown in Table 8, the tolerance values of collinearity diagnostic index of each variable is between 0.740 and 0.878, the value of variance inflation factor (VIF) is 1.139~1.352 < 2, and the P value of t-test of each variable of the equation is less than 0.01, indicating that the parameters in the regression model are reliable and the equation is established. Therefore, the regression model of W's match with the top 11-20 world athletes is as follows:

Y =10.498 + 15.743 * X1 + 14.178 * X2 +19.210 * X3 + 20.561 * X4

The results of standard regression coefficients in Table 8 represented that the influence of the four independent variables on the dependent variables can be determined. Based on this, the influence of the technical and tactical strength difference between W and the top 11-20 world athletes on the CP is as follows: X3 (competition strength in attack after receive phase) > X1 (competition strength in attack after serve phase) > X4 (competition strength in rally phase II) > X2 (competition strength in rally phase I).

Table 8. Results of model coefficient test for W's match with the top 11-20 world athletes

Model	Non-standardized coefficients		Standard coefficient	t	Sig.	Collinearity statistics	
	B	Standard error	Trial version			Tolerance	VIF
(constant)	10.498	0.221		47.409	0.000		
X1	15.743	1.962	0.563	8.025	0.000	0.794	1.260
X2	14.178	2.794	0.369	5.074	0.000	0.740	1.352
X3	19.210	2.363	0.565	8.13	0.000	0.807	1.238
X4	20.561	2.772	0.495	7.417	0.000	0.878	1.139

3.5. Evaluation of W's Strength Difference in the Match with the Top 20 or Above World Athletes

Table 9 shows the evaluation results of W's strength difference in the match with the top 20 world athletes. According to the statistical results of the data in the table, W in 12 matches, the strength difference between the attack-after-serve phase and the rally phase II fell within "large positive disparity" ($\bar{X}$ = 0.017, $\bar{X}$ = -0.005), and the strength difference between the rally phase I and the attack-after-receive phase belongs to "equal strength" ($\bar{X}$ = 0.036, $\bar{X}$ = 0.002). It exhibits that when W competes with athletes with a lower ranking in the world, the technical level in the attack- after-serve phase and the rally phase II were prominent.

From the six matches won by W (No. 1, 4, 5, 8, 11 and 12), he won the match with two German athletes DD and RR (No. 4 and 11) with at least two indicators reaching the level of "large positive disparity". But from the overall point of view of the two matches, W's competitive level was still relatively stable, and none of the indicators is below the level of "large negative disparity". At least one of the other winning matches has reached the level of "wider positive disparity", and the two indicators are above the level of "equal strength". In W's negative matches (No. 2, 3, 6, 7, 9 and 10), his three matches (No. 2, 6 and 9) have at least one index reaching the level of "wide positive disparity" or "large positive disparity" in terms of three indexes, but he was still defeated. This is mainly due to the obvious loss of points in the key techniques (the first four strokes techniques). For example, the loss of points in the technical and tactical play of the attack after serve phase and the attack after receive phase are relatively large. Even in the match with OA (No. 9), the strength difference of rally phase II reaches the

level of "wide positive disparity" and the three indicators are "equal strength"; but W still lost the match. However, it is worth noting that in this match, W lost in a big margin (3:4) and took the lead in a small margin (62/57), which is also the case in the other two matches (No. 3 and 7). This demonstrates that W has a great probability of losing the match in the rally phase when the advantages of key techniques (the first four strokes techniques) are not obvious. It further explains that to win in the match, W must securely take the first four strokes techniques, and strengthen the rally ability on this basis.

Table 9. Evaluation list of strength difference between W and the athletes ranking 20 or above in the world

S/N	Athlete	Attack-after-serve phase		Rally phase I		Attack-after-receive phase		Rally phase II		Results
		SD	Evaluation	SD	Evaluation	SD	Evaluation	SD	Evaluation	
1	JN	0.214	++	0.029	0	0.014	0	0.000	+	4:0(44/26)
2	AN	0.046	+	0.065	+	-0.111	--	-0.018	+	2:4(53/55)
3	UA	0.021	+	-0.014	-	0.014	0	-0.007	+	3:4(74/72)
4	DD	-0.009	0	0.060	+	0.034	+	0.000	+	4:3(64/52)
5	UA	0.034	+	0.080	+	0.080	++	-0.023	0	4:1(51/36)
6	KA	-0.114	--	0.048	+	-0.076	--	0.019	++	2:4(46/59)
7	FS	-0.076	-	0.042	+	0.025	0	0.017	++	3:4(60/59)
8	ZH	0.000	0	0.048	+	0.032	+	0.048	++	4:2(71/55)
9	OA	-0.008	0	0.017	0	-0.008	0	0.042	++	3:4(62/57)
10	OI	-0.035	0	-0.024	-	0.012	0	-0.082	-	1:4(36/49)
11	PS	-0.025	0	0.059	+	0.068	+	-0.034	0	4:2(63/55)
12	CG	0.155	++	0.024	0	-0.060	-	-0.024	0	3:1(46/38)
$\bar{X}$		0.017	+	0.036	0	0.002	0	-0.005	+	

3.6. Multiple Regression Analysis of W's Match with the Top 20 World Athletes

The data shown in Table 10 exhibit that in the regression model of W's match with the top 20 or above world athletes, the complex correlation coefficient R, determination coefficient R^2 and adjustment coefficient R^2 are all above 0.94, indicating that the model is of high fitting degree. And the determination coefficient R^2 value is 0.963, which means that the strength difference of the four

indicators can explain the 96.3% change in the competition performance (CP). Analysis of variance reveals that the F value of the regression model was 45.047, $P < 0.01$, indicating that the regression model is significant.

Table 10. Multiple regression model and analysis of variance results of W's match with the athletes ranking 20 or above in the world

R	R^2	Adjusted R^2	Standard error	Durbin-Watson	F value of regression equation	Sig value of regression equation
0.981	0.963	0.941	0.54	2.772	45.047	0.000

According to the data results shown in Table 11, the tolerance value of collinearity diagnostic index of each variable is between 0.905 and 0.977, and the value of variance inflation factor (VIF) is between 1.024 and 1.105 < 2. In the t-test of each variable of the equation, except that the significance value of X2 (strength difference of rally phase I) is less than 0.05, the significance value of other independent variables is less than 0.01, suggesting that the parameters in the regression model are reliable and the equation is established. Based on this, the regression model of W's match with the athletes ranking 20 or above in the world is as follows:

$$Y = 10.273 + 20.718 * X1 + 18.360 * X2 + 17.479 * X3 + 18.104 * X4$$

The results of the standard regression coefficients in Table 11 exhibit that the influence of the four independent variables on the dependent variables can be determined. On account of this, it can be concluded that the influence of W's technical and tactical strength difference in each phase of the match with athletes ranking 20 or above in the world on the competition performance (CP) is as follows: X1 (strength difference in attack after serve phase) > X3 (strength difference in attack after receive phase) > X4 (strength difference in rally phase II) > X2 (strength difference in the rally phase I).

Table 11. Results of model coefficient test for W's match with the athletes ranking 20 or above in the world

Model	Non-standardized coefficient		Standard coefficient	t	Sig.	Collinearity statistics	
	B	Standard error	Trial version			Tolerance	VIF
(constant)	10.273	0.259		39.709	0.000		
X1	20.718	1.812	0.846	11.436	0.000	0.977	1.024
X2	18.360	5.425	0.259	3.384	0.012	0.915	1.093
X3	17.479	2.879	0.449	6.071	0.001	0.977	1.024

X4	18.104	4.837	0.288	3.743	0.007	0.905	1.105

3.7. Kruskal Wallis Test of W and Athletes with Different Strength Rankings

In order to further test the difference of strength difference between W and athletes with different strength rankings (top 10, top 11-20 and 20 or above in the world), a nonparametric Kruskal Wallis test was employed. As per the test results described in Table 12, the strength difference in W's rally phase II presented a very significant difference in the match with athletes with different strength rankings, $P < 0.01$, and there is no significant difference in the index strength difference of other phases, $P > 0.05$. This further verifies that W's competitive level in rally phase II in the match with athletes with different strength rankings fluctuates obviously.

Table 12. Kruskal Wallis test statistics

	Attack-after-serve phase	rally phase I	Attack-after- receive phase	rally phase II
Chi square	1.101	1.423	0.963	14.154
df	2	2	2	2
Asymptotic significance	0.577	0.491	0.618	0.001

3.8. Advantages and Limitations of this Study

This study has certain advantages in the field of table tennis technical and tactical analysis. First, our study was conducted by watching a large number of match videos to collect four phases of table tennis technical and tactical data, which is more objective than single matches or a few matches as the object of analysis in the past. Secondly, we found that multiple regression analysis combined with competition strength difference evaluation can well evaluate the performance of table tennis athletes' four indexes and the competitive status characteristics of both athletes in a match, which can provide theoretical reference for other technical and tactical ability- dominated net-separating antagonistic sports.

However, there are limitations to this study. First, this study showed the phenomenon of "imbalance between the total score and the win-loss relationship" in the data statistics of the athletes in the match, that is, the winner lost the total score to the opponent and won the match score to the opponent, although some studies proved that the probability of this phenomenon in the match is 5%, this study did not provide a unique analysis of this phenomenon in the data mining. Secondly, this study only evaluated and predicted the competition ability from the four-phase index as the observation index, and did not dig deeply into the competition performance characteristics of the athletes, such as the specific skills and tactics used in the attack after serve phase, standing position, and rotation quality of the serve, etc. Therefore, it is suggested that it can be properly included in future research for analysis. In addition, the psychological characteristics of athletes in competition are also the focus of the current research. This study only evaluated the performance of competitive level from a quantitative perspective. However, the final point of the match often depends more on whether the

athletes have strong psychological qualities, so it is necessary to conduct qualitative observation and evaluation of the psychological characteristics of athletes in competition.

4. Conclusion

The multiple regression model is useful for predicting the effectiveness of the technical and tactical indicators of table tennis athletes in the match, and can clearly reflect the strengths and weaknesses of the technical and tactical indicators, which can be applied to the prediction of the competitive effects of other ball games.

W's performance in terms of technical and tactical aspects varies and fluctuates when competing against athletes of different world rankings. His competitive performance in various indicators during matches against players ranked in the world top 10 is mediocre, and in some stages, it even tends to be below average. His relative standout technical performance is not fully realized. In the match between W and the top 10-20 world athletes, the competitive level performance of the attack-after-serve phase and the attack-after-receive phase is more prominent, and the overall competitive performance is relatively stable. In the match between W and athletes ranked after world top 20, the technical level of the attack-after-serve phase and rally phase II is more prominent, and the competitive performance of other phases is average. The Kruskal-Wallis test results reveal significant fluctuations in the competition performance (CP) during the rally Ⅱ phase, demonstrating highly significant differences ($P < 0.01$). In matches, W has a very low probability of winning when not holding an absolute advantage in key techniques (the first four strokes) - particularly pronounced when facing athletes ranked in the world top 20.

Author contributions

Conceptualization, X C; methodology, HG Y; formal analysis, X C; data processing, X C and YX Z; writing—original draft preparation, X C; writing—review and editing, HG Y; project administration, X C; funding acquisition, N/A. All authors have reviewed and agreed to the published version of the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

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

Declarations

Ethical approval and consent to participate

Consent for publication

N/A.

Competing interests

The authors declare that they have no competing interests.

Funding

This work has been supported by the Economic and Social Development and Philosophy and Social Science Fund of Shangrao City (NO. 24YB23); the Education and Research Program for Young Teachers

in Fujian Province (Science and Technology) (NO. JAT220308); and the Science and Technology Project Plan of Putian City (NO. 2023SZ3001PPTXY15); Putian University Research Project "Intervention and Mechanism Study of Fitness Yoga on Upper Cross Syndrome" (No. 2022056). The acquisition of experimental data from these competition videos complies with the current laws of the country. The author does not have a conflict of interest to declare.

References

[1] Zhang, H., Zhou, Z., & Yang, Q. (2018). Match analyses of table tennis in China: a systematic review. Journal of sports sciences, 36(23), 2663-2674.

[2] Straub, G., & Klein-Soetebier, T. (2017). Analytic and descriptive approaches to systematic match analysis in table tennis: A review. German Journal of Exercise and Sport Research, 47(2), 95-102.

[3] Yin, H., Chen, X., Zhou, Y., Xu, J., & Huang, D. (2023). Contribution quality evaluation of table tennis match by using TOPSIS-RSR method-an empirical study. BMC Sports Science, Medicine and Rehabilitation, 15(1), 132.

[4] Sun, X., Lu, L., & Shi, M. (2023, August). Evaluation of Technical Ability of High-Level Table Tennis Players Based on Multimedia Video. In EAI International Conference, BigIoT-EDU (pp. 171-182). Cham: Springer Nature Switzerland.

[5] Sung, C. Y. (2019). A smart analysis of technical skills of top male table tennis players. Smart Science, 7(4), 231-238.

[6] Fuchs, M., Liu, R., Malagoli Lanzoni, I., Munivrana, G., Straub, G., Tamaki, S., ... & Lames, M. (2018). Table tennis match analysis: a review. Journal of sports sciences, 36(23), 2653-2662.

[7] Zhang, S., Chen, G., Wu, Q., & Li, X. (2023). The Interplay Between Table Tennis Skill Development and Sports Performance: A Comprehensive Review. Pacific International Journal, 6(3), 150-156.

[8] Amalia, E. F., Subarjah, H., & Safari, I. (2017). The Influence of Exercise Motivation and Motor Ability towards the Table Tennis Playing Skills. Health and Physical Education, 1, 295-304.

[9] Zaferanieh, A., Haghighi, A. H., Kakhak, S. A. H., Maleki, A., Cè, E., & Esposito, F. (2021). Effect of ballistic and power training on performance adaptations of élite table tennis players. Sport Sciences for Health, 17, 181-190.

[10] Yin, H., & Espeso, L. A. (2024). Application and Assessment of the Technical Effectiveness of Table Tennis Matches by the Principal Component Analysis (PCA)–Cluster Analysis (CA) Composite Model. Journal of Education and Educational Research, 8(1), 106-112.

[11] Grycan, J., Kołodziej, M., & Bańkosz, Z. (2023). Technical and Tactical Actions of the World's Leading Male Table Tennis Players Between 1970 and 2021. Journal of Sports Science & Medicine, 22(4), 667.

[12] Shi, Z., Jia, Y., Shi, G., Zhang, K., Ji, L., Wang, D., & Wu, Y. (2024). Design of motor skill recognition and hierarchical evaluation system for table tennis players. IEEE Sensors Journal.

[13] Wang, J., Wu, J., Cao, A., Zhou, Z., Zhang, H., & Wu, Y. (2021). Tac-Miner: Visual tactic mining for multiple table tennis matches. IEEE Transactions on Visualization and Computer Graphics, 27(6), 2770-2782.

[14] Wang, L., Zhou, Z., & Zou, Q. (2023). Analysis system for table tennis techniques and tactics using data mining. Soft Computing, 27(19), 14269-14284.

[15] Zhang, H., & Zhou, Z. (2017). An analytical model of the two basic situation strategies in table tennis. International Journal of Performance Analysis in Sport, 17(6), 970-985.

[16] Wu, Y., Lan, J., Shu, X., Ji, C., Zhao, K., Wang, J., & Zhang, H. (2017). iTTVis: Interactive visualization of table tennis data. IEEE transactions on visualization and computer graphics, 24(1), 709-718.

[17] Huang, W., Lu, M., Zeng, Y., Hu, M., & Xiao, Y. (2021). Technical and tactical diagnosis model of table tennis matches based on BP neural network. BMC Sports Science, Medicine and Rehabilitation, 13(1), 54.

[18] Li, W. (2024). Tactical analysis of table tennis video skills based on image fuzzy edge recognition algorithm. IEEE Access.