International Journal of **Music Business Research**

Factors Predicting Singers' Work Efficiency and Singers' Singing Peak

Xuejie Huang, Mei Foong Ang*

Department of Music, Faculty of Human Ecology, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

Abstract: The research focusses on understanding the controls and predicting singers' work efficiency, which is measured by rework times (RwT) in the recording studios. We conducted a comprehensive literature synthesis to identify a set of nine critical factors combining advancements in both acoustic field and recording practice. We measured each of these potential variables while 10 individual singers performed the same original song independently. The pieces of music sentences (n = 380) were analyzed by the multiple linear regression method. We found that: 1. In the recording process, the RwT of a singer are predicted mainly by intensity, singing duration (SinD) and rhythm consistency. 2. SinD mediates the connection between intensity and RwT. 3. A singing peak exists in the performance.

Keywords: *music content • recording • singer • singing peak • work efficiency*

1. Introduction

At the core of the music industry's success lies the production of quality music content (Hull et al., 2011; Saragih, 2017; Saragih et al., 2019). From the audience's standpoint, the continued monetisation of music production is intricately tied to the excellence of the content (Cuadrado et al., 2015). In the pre-digitalisation era, achieving high-quality music production demanded substantial investments in human, material and financial resources (Perrett Jr., 2011). As the number of individual music producers and production teams continues to rise, there is an escalating need for cost-effective music production to facilitate the establishment of young musicians in the market (Paul, 2012). In the current music production landscape, the emphasis on highly efficient processes and cost-saving measures significantly contributes to the sustainable growth of the digital music industry (Kim & Kang, 2022). There is a clear call within the music market for small or medium-sized music producers to present innovative, cost-saving and efficient music production models.

Evidence indicates that individuals in key roles, such as singers, significantly influence the economic development and efficiency of the creative industry, particularly in the field of music (Attanasio et al., 2022; Madudová, 2017; Ruoslahti, 2020). Singers operating within the recording studio serve as the crucial link between content production and publication. Their efficiency during the recording process is critical, as it can effectively reduce costs in the upstream sector of the music industry. Traditionally, discussions in the upstream focus on music content, particularly pre-released songs, often accompanied by debates on innovations driven by new equipment or technology. However, there is a noticeable gap in our understanding of the practical working conditions, with insufficient attention given to enhancing recording efficiency. While the importance of minimising rework time (RwT) (the time that the singer needs to repeat certain elements to align with the criteria set by the music production team) in the recording practice is evident, earlier studies have tended to prioritise theoretical aspects, neglecting this critical operational dimension.

* E-mail: meifoong@upm.edu.my

The assessment of performance and vocal quality is consistently evolving. Eidsheim (2011) emphasised a departure from the limited understanding of music centred on its static components—such as pitch, duration and auditory quality—towards a more comprehensive perspective that considers the dynamic interplay of energies and contexts. Vocal singing analysis extends beyond the mere examination of singing or performance techniques; it encompasses the dimensions of sound. With advancements in acoustic measurement, the contemporary assessment model is bridging the divide between acoustic culture and sound ontology. In addition, the definition of a "beautiful sound" has evolved to encompass not only the voice producer but also the sound receiver. This conceptual shift has motivated Gupta et al. (2018) to propose a holistic approach that combines perceptual and computing parameters (PESnQ) for evaluating vocal dimensions. This integrated model considers both subjective perceptions and objective computational analysis, enriching our understanding of the intricate elements that contribute to the assessment of vocal quality.

The PESnQ contains key variables such as pitch (P), rhythm consistency (R), voice quality (Vqual), pronunciation (Pronun), appropriate vibrato (Vib), volume (Vol) and pitch dynamic range (PDR). Computational analysis of these parameters falls within the domain of music information retrieval (MIR), which aims at developing computational tools for processing, searching, organising and accessing music-related data. However, MIR, as a framework, has been limited in vocal appreciation, as its parameters primarily describe a singer's final work. It often lacks consideration for variables beyond vocal elements, such as those encompassed in the interdisciplinary review by Lerch et al. (2020). In evaluating a singer's true proficiency, the improvisational aspect of musical renditions is often neglected. This limitation hinders our awareness of a singer's RwT dedicated to perfecting each musical phrase or the specific instances in MIR recordings where the singer demonstrates optimal vocal ability.

Rink (2002) proposed a two-fold model, categorising the initial approach as "prescriptive", involving the analysis of acoustic material beforehand. The second, termed "descriptive", delves into the examination of the actual performance process. However, certain factors, such as the dimension of singing RwT, singing duration (SinD) and the nature of studio recordings—deemed as "descriptive" variables—have not been integrated into the existing Music Performance Analysis (MPA) that systematically examines and evaluates a musical performance. Drawing inspiration from the Rink (2002) MPA framework, our exploration extended beyond acoustic parameters to include the improvisational conditions influencing a singer's rendition. Consequently, we advocate for a synergistic approach, combining both MIR and MPA methodologies to predict a singer's efficiency, measured through the lens of RwT. Towards the targets of high efficiency in music production and the establishment of robust criteria for selecting exceptional singers, our study focusses on addressing the following questions:

(1)   How can we accurately predict RwT in the recording process?
(2)   Is there a singing peak for a singer?

2. Method

We proposed the Singing Efficiency Model (Figure 1), which combines performance parameters extracted through MIR with the actual performance process represented by SinD. Owing to the diverse rhythms, syllables and pitch ranges throughout the SinD, singers may not strictly adhere to the musical score. In essence, the singer's performance level fluctuates in response to the varying durations they encounter (Umbert et al., 2015). Figure 1 presents the variables involved in the prediction of RwT. The intensity, R, P, Pronun, PDR, VQual and vibration (Vib) are the acoustic variables, the SinD and RwT are the variables from the music recording practice. Evaluating music performance in a meaningful way faces a challenging issue, since singers show various characteristics in different SinDs (Silas et al., 2023). In the current study, we unexpectedly observed a dramatic decrease in RwT for singers as the SinD changed. Hence, we speculated that there were fluctuations in the intensity of a singer throughout the SinD. Furthermore, to gain knowledge on minimal RwT, we further studied whether there is a singing peak session, defined as the period with the most stability and minimal RwT. The "peak experiences" described by Maslow (1971) and the "epiphany" mentioned by Turner & Newman (2005) both highlight the top moments of the singing state in music performance, i.e. singing peak.

We used a set of computer equipment (including a built-in system of Windows 10) and microphone-relevant equipment such as microphone amplifiers, speakers and headphones. ProTools (Edition 12.5) was the audio
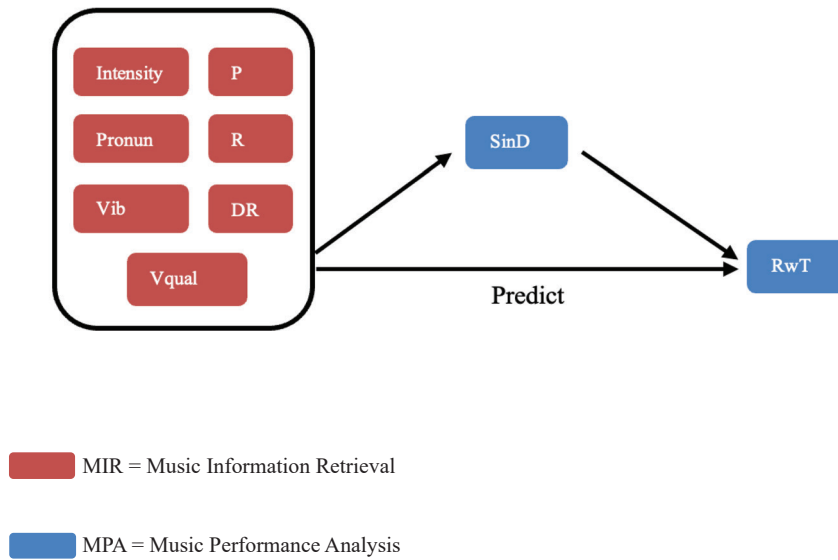
Figure 1. The singing efficiency prediction model. RwT: rework times; SinD: singing duration; DR: dynamic range.

workstation for recording voice waves. Audio signals were collected by the Bluebird Microphone (Baby bottle, USA), and the distance was set at about 15 cm away from the singers' lips. The SSL microphone amplifiers (Super Analog, UK) were used as the signal amplification equipment. The speaker equipment was Adam (S3xh, Germany) and the headphones for the sound monitoring were AKG Headphones (K241, Austria). The audio sample was accomplished at Weake Studio in Chongqing, China, which is an open studio for business that has been acoustically decorated, with a volume of 175 m³ and the reverberation time is about 0.3 s. The recording room (about 5.3 m × 3.2 m × 2.6 m) and the control room (about 8.6 m × 3.2 m × 2.6 m) were divided by three layers of transparent acoustic glasses (2,000 mm × 1,200 mm, Rw = 45 dB).

Ten demo singers coming from the eight provinces of China (Hubei, Chongqing, Beijing, Hunan, Henan, Tipet, Zhejiang and Guangdong) were selected for the recording of an original song. The criteria for the singer selection were a minimum of 3 years' service experience in the demonstration. The song "At the Beginning" (一如最初) was a commercially released composition created by a local composer in Chongqing. All of the recording was completed in Weake Studio in Chongqing. The chosen song was unpublished and unfamiliar to participants, ensuring unbiased vocal performance data. Otherwise, familiarity with a song could compromise the accurate reflection of singing proficiency in the experiment.

The musical style of the selected song is suitable for all vocal groups, making it easier for participants to collect valid samples of their voices. The pitch of the composition goes across two octaves, which allows the dynamic differences of singers to be authentically experienced. The original key of this song is in the key of C♯ major. To ensure the optimal singing process, all vocal sessions were recorded in the singers' comfortable key zones, and the accompaniment of the song was adjusted accordingly for the Tenor (F), Soprano (C♯), Baritone (C♯) and Mezzo-soprano (B♭) parts. Each recording wave was preserved in its dry voice form, a drop of any post-processing, compression or dynamic compression. Each singing wave has been segmented into 38 music sentences according to the musical structure. This approach allowed for a more detailed examination of the acoustic characteristics within the SinD (Figure 2). Moreover, the splitting of the voice is suggested to investigate the internal relationship in the human voice. The similar strategy such as the acoustic study of Han & Zhang (2017), supported successfully observing the voice characteristics.

Figure 2. A qualified wave diagram.

3. Sample Selection

The vocal tracks were initially recorded as dry sounds without compression in the original project. All the recording process aligns with the standard recording industry practice, where the final processed waves are determined by the sound engineer. Consequently, each musical phrase undergoes multiple revisions to meet the industry standards. Notably, the number of rework instances for each musical phrase by each singer has been documented.

The fundamental criterion for selecting the final waves was that both vocal perception and wave feature met the expectations set by the original song. Visually, the waveform should exhibit fullness without any instances of distortion.

4. Measurement

The mean *intensity* for a music sentence was extracted by the Praat (Version 6.3.03, Netherlands). ChiVox was employed to judge and assess the *pronunciation* based on the dimensions of fluency, stress and intonation of morphemes. We used the Melodyne (Editor version, UK) to capture deviations in *rhythm consistency* and *pitch* accuracy. The rhythm consistency was categorised into six levels (1 = messy, 2 = cannot catch, 3 = can catch, but a bit forced, 4 = available, 5 = precise and 6 = strong sense of rhythm). The pitch deviation for each music sentence was recorded, and the deviation value was calculated based on a natural pitch. Sound Meter (Editor version, USA) plugin was used to capture the local *dynamic range* of the singer in each musical phrase. We imported the sound into Praat and utilise the Acoustic Voice Quality Index (AVQI) plugin to assess the corresponding *voice quality*, then directly recorded the scores. The vocal track was input into Praat, and the vibrato extent at the end of each SinD was observed manually. The *vibrato* extent for each music sentence was determined by the difference between the maximum and minimum pitch values. The SPSS (version 27) was used in the statistical analysis. We recorded the number of times the singer rerecords each musical sentence as the *rework times*. Each music sentence labels for a *singing duration* of the singers.

5. Data Analysis

We first checked the correlation between RwT and each predictor variable (R, Intensity, P, DR, Pronun, Vqual and Vib). Given the relatively small sample size entering the model, to balance and compare the goodness of fit among models, we utilised the Akaike Information Criterion (AIC) strategy (Akaike, 1973) for variable selection.

Then, a hierarchical linear regression containing three steps was run. In Step 1, the acoustic variables except intensity were entered into the model. Furthermore, we entered intensity, which shows the highest correlation with RwT, into the model in the second step to study the effect of intensity. Finally, in Step 3, SinD and intensity were both entered into the model, to understand whether SinD mediates the connection between intensity and RwT. In the PROCESS analysis (Na & Hipertensiva, 2022), the mediation effect size of SinD was estimated. Before the analyses, continuous predictors underwent mean-centring. To avoid the potential multicollinearity, an initial analysis was conducted. The examination revealed tolerance rates of .76 and .85, and a variance inflation factor (VIF) of 1.17 and 1.32 for SinD and Intensity. These outcomes suggest the absence of any multicollinearity issues (O'Brien, 2007).

Finally, the values meeting efficient and optimal vocal states were selected, and frequency statistics were conducted. To better examine whether there is a singing peak, the song was categorised into different parts, namely verse1, chorus1, verse2, chorus2, post-chorus and outro. This systematic approach allows for a more thorough analysis of vocal performance across distinct sections of the song.

6. Results

As can be seen in Table 1, most variables showed a significant correlation in the singing data, which supported the previous studies in vocal production (Banse & Scherer, 1996). It presents that there's a significant and negative correlation between R ($r = -0.398$; $p < 0.01$, Table 1) and RwT, SinD ($r = -0.384$; $p < 0.01$) and RwT, Intensity

	R	SinD	Intensity	P	Pronun	DR	Vqual	Vib
RwT	−0.398**	−0.384**	−0.563**	0.181**	−0.116*	−0.230**	−0.059	−0.152**
R		0.148**	0.319**	−0.092	0.027	0.178**	0.003	−0.005
SinD			0.193**	0.261**	0.103*	0.429**	−0.098	0.101*
Intensity				0.139**	0.045	0.199**	0.009	0.093
P					−0.041	−0.109*	0.033	−0.069
Pronoun						0.019	−0.023	0.125*
DR							−0.093	0.184**
Vqual								0.058

P, pitch; Pronun: Pronunciation; R, rhythm consistency; RwT: rework times; SinD: singing duration; Vqual, voice quality.
*$p < 0.05$; **$p < 0.01$.

Table 1. RwT correlated with the acoustic variables.

($r = −0.563$; $p < 0.01$) and RwT, DR ($r = −0.23$; $p < 0.01$) and RwT, Vib ($r = −0.152$; $p < 0.01$) and RwT. A significant positive correlation is found between P ($r = 0.181$; $p < 0.01$). There is a significant and negative low correlation between Pronun ($r = −0.116$; $p < 0.05$) and RwT.

All the variables were entered and AIC strategy (Akaike, 1973) was employed for variable selection. In model selection, models with smaller AIC values are considered for adoption (Portet, 2020). The lowest values (AIC = 1163.11) in a model that included the R, SinD, Intensity, P, Pronun, DR, Vib, and VQual were selected. To examine whether the mediating effect of SinD is established, we conducted a hierarchical regression.

In Step 1, as shown in Table 2, the model incorporating all acoustic variables, excluding intensity, yielded an R^2 of 24% ($p < 0.001$). Upon introducing intensity, the R^2 experienced a significant boost to 40% ($p < 0.001$). Further, by incorporating both SinD and intensity, the R^2 demonstrated a notable increase to 45% ($p < 0.001$). In Step 3, the contribution of SinD and intensity to the prediction of RwT is highly significant.

As seen in Table 2, after controlling for R, P, DR, Pronun and Vib, those with higher intensity reported fewer RwT ($\beta = −0.44$, $p < 0.001$). Moreover, those with higher intensity through SinD reported fewer RwT ($\beta = −0.27$, $p < 0.001$). The main effects in Step 2 explained an additional 16% ($\Delta R^2 = 0.40$, $p < 0.001$) of the variance in RwT. Step 3 explained an additional 5% ($\Delta R^2 = 0.45$, $p < 0.001$) of the variance in RwT. Therefore, the mediation effect of SinD for the connection between intensity and RwT is significant ($\beta = −0.54$, $p < 0.01$). To further investigate the interaction between intensity and RwT, the macro-PROCESS 4.1 (Model 4) was run.

As depicted in Figure 1, the dependent variable (Y) was the RwT. The potential mediators were the SinD (M) and the independent variable (X) was the intensity. The three model steps confirmed the mediation process. Initially, the Intensity on the SinD model confirmed a significant relationship between SinD and intensity ($R^2 = 0.036$, $p < 0.001$), shown in Table 3. Subsequently, the results indicated significance for both intensity and SinD in the RwT model ($R^2 = 0.39$, $p < 0.001$). Finally, the intensity on the RwT model was also found to be significant ($R^2 = 0.31$ $p < 0.001$). This shows that the proposed mediators, particularly SinD (M), likely played a mediating role in the influence of intensity on RwT.

The estimated indirect effect of intensity on RwT mediated by SinD was calculated to be significant (coefficient = −1.06, 95% confidence interval [CI] [−1.22, −0.90], Table 4). The estimated direct effect of intensity on RwT mediated by SinD was significant (coefficient = −95, 95% CI [−1.11, −0.80]). A bias-corrected bootstrapping method was used to compute the value of the indirect effect, which indicated that the mediating effect was significantly different from 0 at $p < 0.05$, as the 95% CI did not contain zero. Therefore, the results revealed that SinD (M) mediated the effect of intensity on RwT.

In investigating the presence of a singing peak, we examined the relationship between SinD and the mean RwT of the singers. The overarching pattern observed indicated a decline in RwT as SinD increased. Notably, from the 23rd to the 35th duration, the singer consistently experienced rework occurrences ranging between 0 and 2, implying a period of high efficiency.

To delve deeper into the alterations in sound characteristics across different segments of the song, we aligned the SinD with the musical structure. Specifically, we categorised the durations into various sections: verse 1 (duration 1–8; n = 80), chorus 1 (duration 9–16; n = 80), verse 2 (duration 17–20; n = 40), chorus 2 (duration 21–28; n = 80), post-chorus (duration 29–36; n = 80) and outro (duration 36–36; n = 20).

Model Variables	B	S.E		95% CI		AIC	R^2	ΔR^2
				LL	UL			
Step 1						1,633.49	0.24***	0.23
R	−0.601	0.097	−0.27***	−0.075	−0.039			
P	0.012	0.007	0.072**	−0.936	−0.626			
DR	−0.038	0.017	−0.09**	−0.010	0.016			
Pronun	−0.168	0.110	−0.06	−0.032	0.035			
Vib	−0.001	0.000	−0.053	−0.429	0.099			
Vqual	−0.087	0.140	−0.025	−0.001	0.000			
Step 2						1,542.07	0.40***	0.40
Intensity	−0.826	0.082	−0.44***	−0.988	−0.664			
Step 3						1,163.11	0.45***	0.45
SinD	−0.057	0.009	−0.27***	−0.075	−0.039			

AIC, Akaike Information Criterion; CI, confidence interval; LL, lower limit; Pronun: Pronunciation; S.E., Standard Error; UL, upper limit.
**$p < 0.01$; ***$p < 0.001$.

Table 2. Linear regression analysis examining the relationship between predictive variables and SinD.

Outcome variable	Predictive variable	R^2	S.E	F	95% CI	
					LL	UL
SinD	Intensity	0.036***	0.46	14.15	0.82	2.61
RwT	Intensity		0.08	121.43	−1.11	−0.80
	SinD	0.39***	0.01		−1.08	−0.04
RwT	Intensity	0.31***	0.08	169.31	−1.22	−0.90

CI, confidence interval; LL, lower limit; RwT: rework times; S.E., Standard Error; SinD: singing duration; UL, upper limit.
Note: ***$p < 0.001$. n = 380.

Table 3. Mediation effect test of SinD.

Variable	Effect type	Coefficient	S.E	95% CI	
				LL	UL
SinD	Total effect	−1.06***	0.081	−1.22	−0.90
	Indirect effect	−0.11***	0.034	−0.18	−0.04
	Direct effect	−0.95***	0.078	−1.11	−0.80

CI, confidence interval; LL, lower limit; S.E., Standard Error; SinD: singing duration; UL, upper limit.
Note: ***$p < 0.001$. n = 380.

Table 4. Mediation effect size of SinD.

In defining the singing peak as a state characterised by consistently optimal performance across various parameters, we conducted frequency statistics on the best-value performance. The criterion for rhythm selection was set at the maximum level of 6. Concerning RwT, a value of zero was considered optimal. While literature lacked a definitive intensity value for the singing peak, previous research suggested intensity values ranging from 65 dB to 102 dB in professional singers (Titze, 1992; Scherer et al., 2017b; Petekkaya et al., 2018). After conducting an independent samples t-test, we found that 75 dB is the most suitable standard for all the singers in this study. Hence, the intensity value which was 75 dB was selected. Regarding pitch accuracy in the recording, experienced musical practitioners recommended using a criterion of 50 cents (ct) (Hutchins et al., 2014) to assess the good pitch control of singing. Therefore, deviations in pitch accuracy ≤50 ct were set as the criterion.

Firstly, the singing without rework occurs most frequently during chorus 2 and post-chorus section. The best value of intensity occurred 43 times in chorus 2, accounting for 53.75%, and 40 times in post-chorus, accounting for 50.0%. Meanwhile, the percentages of instances with no rework in chorus 2 and post-chorus are 50.0% and 47.5%,

respectively, making them the highest among all durations with zero rework. Although the best values for rhythm frequently appear in chorus 1 (68.75%), this section performs poorly in terms of RwT (13.75%). Therefore, chorus 1 is not considered to be part of the best singing states. The pitch deviation of fewer than 50 cents accounts for 100% in chorus 2 and 98.75% in post-chorus. In summary, the singer demonstrates the best vocal performance in chorus 2 and post-chorus, representing the peak of singing. The outcome is consistent with Hypothesis 3.

7. Discussion

Intensity, as a multidimensional variable, is often confused with the concepts of volume and loudness. Intensity is more applicable to explaining a singer's vocal mechanism, including lung pressure and airflow speed. Our study initially validates that excellent sound is associated with outstanding intensity, and this association is consistent. During the period of optimal intensity, there is also a relative improvement in work efficiency.

7.1. Recording Strategy
The SinD and its mediating effect can be strategically leveraged in the recording industry. Notably, the findings highlight that singers achieve their best intensity and rhythmic control during the chorus. This insight suggests innovative strategies for recording practices. For instance, initiating the recording process from the chorus offers a swift approach for singers to attain their optimal singing state. Recording voices in this unconventional order may facilitate easier management of intensity and rhythmic consistency.

7.2. Music Training Strategy
The finding of the singing peak indicates that variables such as pitch, rhythm and intensity have reached the best state within the scope of the ability of a singer. "Singing peak" refers to the consecutive points where the vocal parameters reach their optimal values, ensuring the highest efficiency. In the context of vocal training, it indicates the singer's optimal state during singing, which has been proven to be facilitated through methods such as vocal warm-up exercises (Grady & Cook-Cunningham, 2020). The relationship between the appearance of the singing peak and acoustic parameters indicates the better the intensity and rhythm consistency a singer is at, the more probability that the singing peak will come.

In the field of vocal listening, the traditional determination of a good voice relies on perception. Surprisingly, singers or listeners may not have perceived them when the singing state reaches its peak. For instance, the audience values a singer based on a particular piece, yet when they sing other compositions, they may not be able to control their voice as effectively as before. In the model we have established, the emphasis is on the potential of the voice, meaning it possesses high productivity while maintaining good quality. This model avoids misjudging a high-potential voice simply because the singer may not be familiar with a particular music piece. In this condition, the singing efficiency model could provide a more scientific approach to examining the singer's level. Especially in the limited budget project, the prediction model is proposed to be employed in rapidly selecting the good singers. Moreover, an approach utilising both the singing efficiency model and perception is suggested to be employed during auditions at a music conservatory, so that a comprehensive selection of high-potential voices will be achieved.

8. Implications

Intensity studies are often associated with emotion recognition, and there is considerable research on how intensity can be used to identify emotions (Alku et al., 2002; Chen et al., 2012; Garcia-Garcia et al., 2017; Wilson, 2021). The analysis of intensity in emotion recognition is currently confined to identifying a few basic emotions and does not reveal whether a singer possesses consistent emotional expression. Further identification of emotion was suggested to focus on observing intensity clues rather than specific points. It is possible that singers exhibit a high degree of emotional coherence during the singing peak. If this aspect is validated, our further research may potentially offer a method to identify emotional coherence using intensity lines.

In our observations, it was noted that certain singers (S1, S5, S8 and S9) consistently achieved optimal values for intensity and rhythm consistency, both in the chorus and post-chorus sections. This suggests a variability among

individuals in their ability to enter and sustain a singing peak. To delve deeper into this phenomenon, especially across different demographics, a comprehensive analysis such as cluster analysis is highly recommended for a more nuanced discussion.

Furthermore, our observations revealed that singers exhibited behaviours indicative of satisfaction, excitement, increased verbal expression and a desire for communication when they reached their singing peak. Therefore, a fascinating avenue for exploration would involve delving into the associated behaviours of the singing peak. This additional layer of analysis could provide valuable insights into the intensity and communicative aspects linked to optimal singing performance.

9. Limitations

The research has certain limitations. The sample group is confined to professional singers, consequently, the context of a "singing peak" is also confined to voices that have received professional training.

Given that close collaboration is a prerequisite for musicians to work effectively in the recording studio (Herbst & Tim, 2018), there were instances of reworks caused by cooperation and communication issues among studio musicians. However, in this research, the model estimated factors based on acoustic parameters, explaining half of the R^2 value in regression. The remaining half, likely attributed to reworks resulting from communication issues, was not considered in the model. To enhance the efficiency of singers in the recording industry, it is crucial to broaden the perspective beyond technical aspects in further study.

From the curve of RwT (Figure 3) and the frequency statistics of vocal parameters (Table 5), it is evident that singers cannot sustain the singing peak for a long time. From this perspective, the remaining singing peak is limited
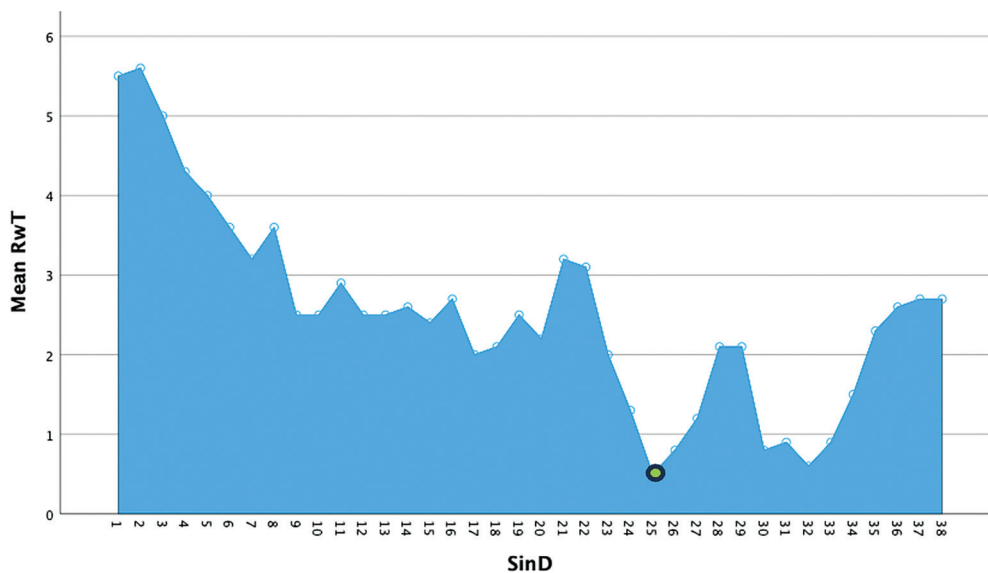

Figure 3. Mean RwT changed by SinD. RwT: rework times; SinD: singing duration.

Duration	Variable			
	RwT = 0	Intensity ^{3} 75	R = 6	P £ 50
Verse1	12 (15.0%)	26 (32.5%)	30 (37.5%)	77 (96.25%)
Chorus1	11 (13.75%)	28.75 (35.94%)	55 (68.75%)	78 (97.5%)
Verse 2	13 (16.25%)	11 (27.5%)	21 (52.5%)	38 (95.0%)
Chorus 2	40 (50.0%)	43 (53.75%)	51 (63.75%)	80 (100%)
Post-chorus	38 (47.5%)	45 (56.25%)	58 (72.5%)	79 (98.75%)
Outro	6 (30.0%)	8 (40.0%)	10 (25.0%)	18 (90.0%)

Table 5. Frequency of the acoustic best performance in each part of the song.

to be observed in this selected song. In cases where a song features an extended chorus, there will be a scholarly interest in exploring whether the duration of the singing peak is proportionally prolonged, warranting a thorough investigation.

10. Conclusion

During the recording process, a singer's RwT can be predicted mainly by intensity, SinD and R. Notably, there is a significant negative correlation between SinD and RwT. SinD serves as a mediating factor in the relationship between intensity and RwT. Furthermore, the singers exhibit singing peaks in specific sections, showcasing exceptional high-pitch control, sustained high vocal intensity and consistent rhythm during high-efficiency work.

Acknowledgements

References

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. *Foundations and Basic Theory*, *1*, 267–281.

Banse, R., & Scherer, K. R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

Chen, X., Yang, J., Gan, S., & Yang, Y. (2012). The contribution of sound intensity in vocal emotion perception: Behavioral and electrophysiological evidence. *PLoS ONE*, *7*(1). https://doi.org/10.1371/journal.pone.0030278

Cuadrado, M., Miquel, M. J., & Montoro, J. D. (2015). Analyzing Music Services Positioning through Qualitative Research. *Market-Tržište*, *27*(2), 251–263. https://doi.org/https://hrcak.srce.hr/151325

Eidsheim, N. S. (2011). Sensing Voice-Materiality and the Lived Body in Singing and Listening.pdf. *Science&Society*, *6*(2), 133–155.

Grady, M. L., & Cook-Cunningham, S. L. (2020). The Effects of Three Physical and Vocal Warm-up Procedures on Acoustic and Perceptual Measures of Choral Sound: Study Replication With Younger Populations. *Journal of Voice*, *34*(4), 647.e15-647.e22. https://doi.org/10.1016/j.jvoice.2018.12.009

Han, Q., & Zhang, R. (2017). Acoustic Analyses of the Singing Vibrato in Traditional Peking Opera. *Journal of Voice*, *31*(4), 511.e1-511.e9. https://doi.org/10.1016/j.jvoice.2016.11.016

Herbst, J., & Tim, A. (2018). The Work Realities of Professional Studio Musicians in the German Popular Music Recording Industry: Careers, Practices and Economic Situations. *IASPM Journal*, *8*(2), 18–37. https://doi.org/10.5429/2079-3871(2018)v8i2.3en

Hutchins, S., Larrouy-maestri, P., & Peretz, I. (2014). Singing ability is rooted in vocal-motor control of pitch. *Atten Percept Psychophys*, *July*, 2522–2530. https://doi.org/10.3758/s13414-014-0732-1

Kim, J., & Kang, M. Y. (2022). Sustainable success in the music industry: Empirical analysis of music preferences. *Journal of Business Research*, *142*(January), 1068–1076. https://doi.org/10.1016/j.jbusres.2022.01.021

Lerch, A., Arthur, C., Pati, A., & Gururani, S. (2020). An Interdisciplinary Review of Music Performance Analysis. *Transactions of the International Society for Music Information Retrieval*, *3*(1), 221–245. https://doi.org/10.5334/tismir.53

Madudová, E. (2017). Creative industries value chain: The value chain logic in supply chain relationships. *Marketing and Branding Research*, *4*(3), 227–235. https://doi.org/10.33844/mbr.2017.60236

Maslow, A. (1971). Farther reaches of human nature.

Na, D. E. C., & Hipertensiva, C. (2022). Introduction to Mediation, Moderation and Conditional Process Analysis-A Regression-based Approach.pdf.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Paul, D. (2012). A musicianship for the recording studio. In *Sound musicianship: Understanding the crafts of music* (pp. 1–6). Cambridge Scholars.

Perrett Jr., H. H. (2011). New Business Models for Music. *Villanova Sports and Entertainment Law Journal*, *18*(January), 63.

Petekkaya, E., Özandaç Polat, S., & Hilmi Yücel, A. (2018). Comparative Analysis of Vocal Intensity and Aerodynamic Parameters between Singers and Non-Singers. *SM Otolaryngology*, *2*(1), 1–4. https://doi.org/10.36876/smotol.1016

Portet, S. (2020). A primer on model selection using the Akaike Information Criterion. *Infectious Disease Modelling*, *5*, 111–128. https://doi.org/10.1016/j.idm.2019.12.010

Rink, J. (2002). Analysis and (or?) performance. In J. Rink (Ed.), *Musical Performance: A Guide to Understanding* (pp. 35–58). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511811739.004

Ruoslahti, H. (2020). Complexity in project co-creation of knowledge for innovation. *Journal of Innovation and Knowledge*, *5*(4), 228–235. https://doi.org/10.1016/j.jik.2019.12.004

Saragih, H. S. (2017). Critical success factors of recording studio in Indonesia. *Asia Pacific Journal of Marketing and Logistics*, *29*(3), 686–699. https://doi.org/10.1108/APJML-07-2016-0123

Saragih, H. S., Simatupang, T. M., & Sunitiyoso, Y. (2019). Co-innovation processes in the music business. *Heliyon*, *5*(4). https://doi.org/10.1016/j.heliyon.2019.e01540

Scherer, K. R., Sundberg, J., Fantini, B., Trznadel, S., & Eyben, F. (2017). The expression of emotion in the singing voice: Acoustic patterns in vocal performance. *The Journal of the Acoustical Society of America*, *142*(4), 1805–1815. https://doi.org/10.1121/1.5002886

Silas, S., Müllensiefen, D., & Kopiez, R. (2023). Singing Ability Assessment: Development and validation of a singing test based on item response theory and a general open-source software environment for singing data. *Behavior Research Methods*, *0123456789*. https://doi.org/10.3758/s13428-023-02188-0

Titze, I. R. (1992). Vocal intensity in speakers and singers. *Journal of the Acoustical Society of America*, *91*(5), 2936–2946. https://doi.org/10.1121/1.402929

Turner, V., & Newman, F. (2005). Being Who You Aren't ; Doing What You Can't: Community music therapy & the para-doxes of performance. *Voices: A World Forum for Music Therapy*. https://doi.org/doi.org/10.15845/voices. v5i3.229

Umbert, M., Bonada, J., Goto, M., Nakano, T., & Sundberg, J. (2015). Expression Control in Singing Voice Synthesis. *IEEE Signal Processing Magazine*, *32*(6), 55–73. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7298564

Wilson, H. (2021). Measuring the Association Between Vocal Effort and Intensity with Emotion Picture Viewing using the Borg Centimax.