Bridging the Rhythmic Gap: A User-Centric Approach to Beat Tracking in Challenging Music Signals

António Sá Pinto[1,2] and Gilberto Bernardes[1,2]

[1] Faculdade de Engenharia da Universidade do Porto, Porto, Portugal
[2] INESC TEC, Porto, Portugal
asapinto@fe.up.pt

Abstract. Deep-learning beat-tracking algorithms have made significant advancements in recent years. However, despite these advancements, challenges persist when processing complex musical examples, which are often under-represented in training corpora. Expanding on our prior work, this paper delves into our user-centric beat tracking approach by subjecting it to highly challenging musical pieces. We probe the adaptability and resilience of our methodology, illustrating its ease of integration with state-of-the-art techniques through minimal user annotations.
The chosen samples, namely, Uruguayan *Candombe*, Colombian *Bambuco*, and Steve Reich's *Piano Phase*, not only demonstrate our method's efficacy but also exemplify challenging rhythmic dissonance effects such as *polyrhythms*, *polymetres*, and *polytempi*. Thereby, we demonstrate the applicability of our human-in-the-loop strategy in the domain of Computational Ethnomusicology, confronting the prevalent Music Information Retrieval (MIR) constraints found in non-Western musical scenarios. Our approach enables notable improvements in terms of the F-measure, ranging from 2 to 5 times the current state-of-the-art performance. In terms of the annotation workflow, these results translate into a minimum reduction of 50% in the number of manual operations required to correct the beat-tracking estimates.
Beyond beat tracking and computational rhythm analysis, this user-driven adaptation suggests wider implications for various MIR technologies, particularly when music signal ambiguity and human subjectivity challenge conventional algorithms.

Keywords: User-Centred, Transfer Learning, Beat Tracking, Computational Ethnomusicology

1 Introduction

Rhythm is a fundamental aspect of music, making computational rhythm analysis a critical topic within Music Information Retrieval (MIR). This area involves tasks such as tempo determination, rhythmic pattern recognition, and metre determination [9]. Algorithmic beat tracking, the automatic detection of a musical signal's pulse, plays an

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

essential role in various MIR applications that require the parsing of musical *time*, i.e., the beat. In the past decade, beat tracking has seen significant progress, with the current state-of-the-art achieving accuracy levels over 90% on benchmark datasets [5, 2]. However, even these advanced methods can face challenges with complex rhythms, especially if they differ from the features of their training data. These challenges are amplified in specialised areas like Computational Ethnomusicology (CE) [20]. In this domain, the availability of annotated datasets is limited, and the need for specalised cultural knowledge to annotate unique rhythmic examples is crucial. Due to these limitations, many musical traditions remain under-represented in MIR research. This gap highlights a known issue in MIR systems: a primary focus on Western (or *Eurogenetic*) music at the expense of diverse global genres and expressions [4, 8].

To overcome these obstacles, adaptive methods have been proposed for tasks like beat tracking [7] and metre determination [19]. While genre-aware knowledge models might provide solutions, they lack scalability. Fiocchi et al. [6] explored how beat tracking knowledge transfers from mainstream Western to Greek music, but their approach, besides being computationally intensive, did not perform as well as training on the same dataset from scratch and yielded less than satisfactory results on the established *SMC* dataset [10], designed with a focus on challenging musical audio examples.

In light of these shortcomings, we shifted towards a more streamlined solution. Our approach harnesses minimal user annotations to optimise a state-of-the-art beat tracker. In earlier works [16, 15], we introduced this user-centric method, aiming for very high accuracy on specific music pieces. Designed for computational efficiency and compatibility with personal computing devices, our methodology has outperformed established methods across various datasets, most notably on the demanding *SMC* dataset [14].

In this study, we expand the scope of our approach beyond Western music. We evaluate our beat-tracking method using challenging datasets such as the Uruguayan *Candombe* and the Colombian *Bambuco*, both distinguished by their respective *polyrhythm* and *polymetre* features. These musical traditions, with their intricate rhythmic structures, serve as a rigorous test bed to assess the adaptability and robustness of our method. Moreover, we apply our technique to Steve Reich's *Piano Phase*, a composition renowned for its innovative use of concurrent *tempi*. The choice to analyse this piece subjects our method to a formidable challenge: to our knowledge, it is the first reported attempt at beat tracking a *polytempo* composition. Our findings indicate that our method effectively manages diverse rhythmic intricacies, allowing for the streamlined adaptation of a leading beat-tracker across a spectrum of musical styles and genres.

2 Rhythmic Dissonance Challenges

Rhythm serves as a foundational scaffold for many musical traditions. Particularly, within African heritage cultures, there is a notable use of complex rhythmic techniques such as *polyrhythms*, *polymetres*, and to a lesser extent, *polytempi* [1]. While these rhythmic intricacies contribute to the distinctiveness of these traditions, they introduce unique challenges in Music Information Retrieval (MIR). In this section, we briefly address the concept of rhythmic dissonance, emphasizing its manifestations in the datasets selected for our study.

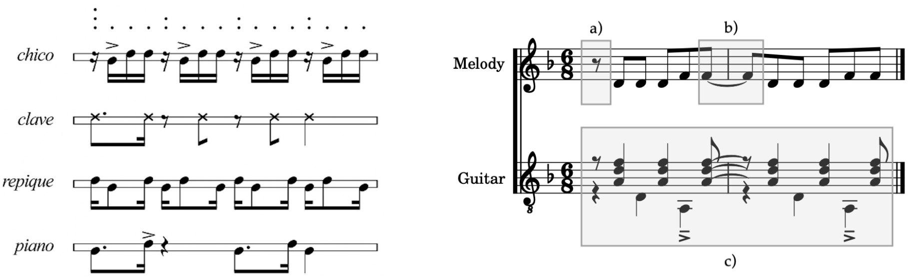

Fig. 1: *Left*: Interaction of the main *Candombe* patterns and the resulting metric structure levels (adapted from [13]). *Right*: Colombian *Bambuco* pattern showing a) downbeat in a rest; b) caudal syncopation; and c) guitar pattern suggesting 6/8 at the top voice and 3/4 at the bass voice (adapted from [3]).

Polyrhythm in Uruguayan *Candombe*: *Candombe* is an African-origin rhythm prominent in Uruguay and, to a lesser extent, in other South American countries [18]. Musically, as illustrated in Fig. 1, it is characterised by the interplay of three percussion instruments: the *chico*, the *repique*, and the *piano*, with an additional time-line pattern called *clave*, shared by the three drums [11]. This combination produces a typical rhythmic structure consisting of a four-beat measure evenly divided into 16 tatums, typically played at a tempo of about 110–150 bpm. *Candombe* distinguishes itself from other rhythms through two features that connect it to Afro-Atlantic music traditions [13]: a) the pulse pattern emphasises the second tatum rather than the one on the beat, and b) the clave divides the 16-tatum cycle irregularly (3+3+4+2+4), with only two of its five strokes synchronised with the beat. This interplay creates an overall polyrhitmic texture. Moreover, in actual performances, the primary pattern of *repique* leans towards a triplet feeling, and although the *chico* drum establishes the metrical foundation, its pattern exhibits a contraction of inter-onset intervals (IOIs). These unique characteristics of *Candombe* present challenges for both untrained listeners and standard beat-tracking algorithms, making it a challenging test case for evaluating our user-driven approach.

Polymetre in Colombian *Bambuco*: *Bambuco* is a Colombian traditional music genre known for its rhythmic complexity, characterised by heavy syncopation, odd accents, and a certain degree of rhythmic freedom, including tempo variations and micro-timing [3]. Its most distinctive aspect is the polymetric nature, resulting from the superposition or alternation of musical elements in two metres: a simple metre (3/4) and a compound one (6/8), as illustrated on the right part of Fig. 1. This phenomenon, commonly known as "hemiola" or the equivalent Latin term "sesquialtera", is relatively common in other South American musical genres [18] but poses a challenge for computational metre and beat-tracking analysis of *Bambuco*. As illustrated by the guitar voice, depending on the simple or compound metre interpretation, the beats' locations do not align, except for the downbeat. This indicates a close relationship between the tasks of metre analysis and beat tracking. Essentially, it implies that we can deduce the metric interpretation from the placement of the beats. These properties make *Bambuco* an ideal test case. More specifically, while our approach primarily targets beat tracking, it also informs metre analysis due to the interconnected nature of these rhythmical facets.

Polytempo in Steve Reich *Piano Phase*: Steve Reich's *Piano Phase* stands out as an interesting example of *polytempo*, a phenomenon mostly absent from mainstream music genres and unrepresented in datasets used to train deep-learning beat-tracking models. This rhythmic dissonance effect presents a significant challenge for general-purpose beat-trackers, as it involves concurrent and isochronous pulses within the same music piece. This compositional technique is primarily found in avant-garde Western music, with Charles Ives's *Symphony no. 4* being considered the earliest formalised work featuring *polytempo*. Later, composers such as Conlon Nancarrow or György Ligeti explored this approach. Steve Reich's *phasing* is a unique manifestation of *polytempo*, where identical phrases are played simultaneously at slightly different *tempi*, creating a gradual phase shift. *Piano Phase* brings Reich's technique to live performance (a rendition of the original score is shown in Fig. 2), complete with a detailed set of instructions for performance, which we briefly summarise:

1. *One performer starts, the other fades in unison (bars* 1–2*), and both continue playing the pattern over and over again*;
2. *The first performer keeps a constant tempo. The other performer gradually increases his tempo, until he is one note ahead of the first performer (bar* 3*)*;
3. *After playing in synchronisation for a while, the second performer again begins increasing his tempo, and the phase shifting process starts again (bars* 3-4*)*;
4. *In the first part of the piece, this procedure is repeated twelve times.*

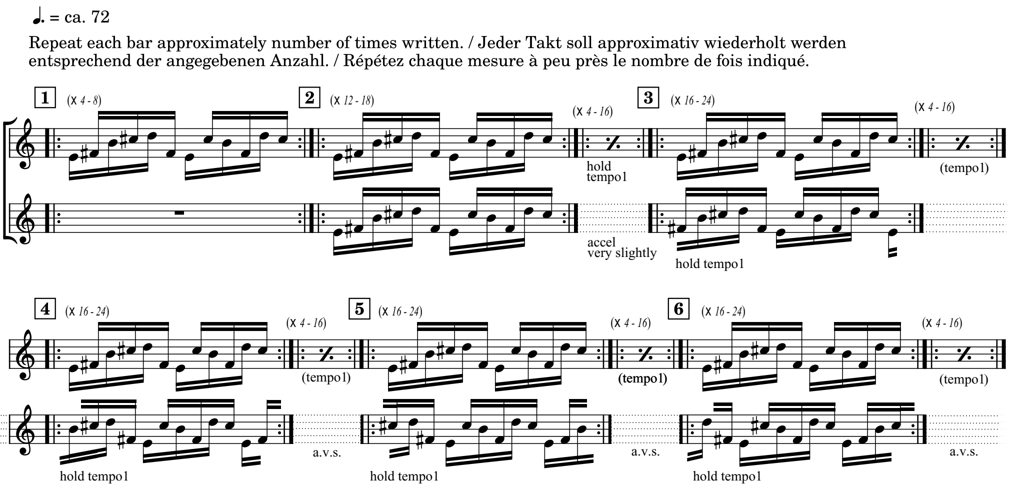

Fig. 2: *Piano Phase*: Partial Reproduction of the Original Score.

3 Methodology

Building on our earlier contributions [14, 15], our approach integrates user knowledge with a state-of-the-art beat tracker [2], enabling direct, content-specific adaptation. Through minimal manual annotation, we tailor this system to the unique characteristics of a musical piece and the user's own subjective musical judgement.

Retraining and Inference: To ensure this paper stands as a self-contained resource, we provide a concise overview of our fine-tuning parameterisation process. For an in-depth understanding and further details on the fine-tuning process, readers are directed to consult [15].

Fine-tuning is allowed for all layers of the baseline network. Given the present task is beat-tracking, the losses for tempo and downbeat tasks on the underlying multitask network [2] are masked. Common practice in transfer learning is followed, thus reducing the learning rate to one fifth of the rate used in the base training. To control network adaptation, we divide the fine-tuning segment into two adjacent, disjoint regions for (re)training and validation, setting a maximum of 50 epochs, and employing learning rate reduction and early stopping strategies. To account for the limited data in the fine-tuning region, target widening and data augmentation are employed. The user-annotated region also serves to parameterise the post-processing Dynamic Bayesian Network (DBN), which extracts beat positions from the Temporal Convolutional Network's likelihood output. For DBN parameterisation, we employ two strategies: 1) adjusting the transition-λ parameter for the adaptive processor type (`pt`), and 2) setting a tempo tolerance window using user annotations as the tempo guide (`tg`). While fine-tuning (`ft`) and data augmentation (`da`) are general user-driven techniques, strategies like the adaptive processor type and tempo guide are specific to networks employing DBN.

Lastly, the length and characteristics of the annotated region, determined by end-users in real-world situations, play a pivotal role in affecting the final performance. In the current experiment, we opted for a relative length, specifically a quarter of the total file length, to standardise the influence of the fine-tuning region length on the evaluation results.

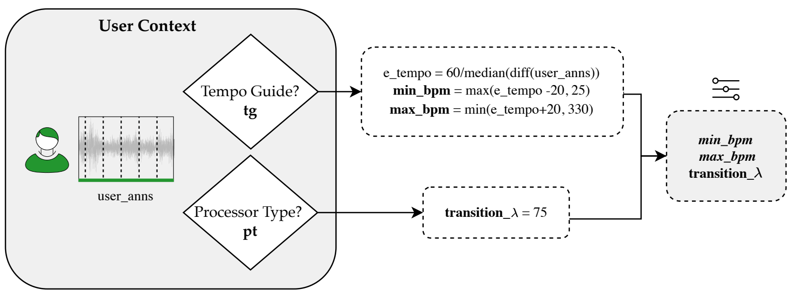

Fig. 3: DBN parameterisation (defaults to `min_bpm`:50, `max_bpm`:215, `transition_`λ:100).

Scope of Evaluation: In this study, we report results with (`fullRes`) and without (`testRes`) the fine-tuned part of the input signal for evaluation purposes, and consider the main combinations of user-driven techniques: fine-tuning (`ft`), data augmentation (da), and DBN customisations (tg and pt). To minimise variability, we adapt the data augmentation procedure from [14] to a deterministic sampling approach based on a linear distribution with a $\pm 30\%$ deviation from the local tempo, calculated using the median inter-beat interval across the annotated region. Results are averaged over three iterations, except for the *Piano Phase* analysis, which results include a single run. While there are 11 combinations of user-driven beat-tracking configurations, this report centres on the primary configurations: ft+da, ft+da+pt, ft+da+tg, and ft+da+tg+pt. These are compared with the state-of-the-art, denoted as baseline (bsl). Occasionally, we reference results from configurations that highlight the standalone application of specific techniques, namely ft, pt, and tg.

Evaluation Metrics: In the present study, we employ both the standard F-measure and a previously proposed *annotation efficiency* (Ae) metric [17] for beat tracking evaluation. The Ae conceptualises beat tracking evaluation from a user workflow perspective, framing it in terms of the effort necessary to modify a series of detected beats to align with the ground-truth annotations. It provides a more intuitive understanding of the evaluation process and aligns better with practical annotation workflows. This is quantified by counting the number of edit operations, specifically insertions and deletions, but also - contrarily to the F-measure -, including the *shifting* of poorly localised individual beats, a very common operation in annotation workflows. This dual evaluation framework, combining both traditional and user-centric metrics, offers a more comprehensive insight into beat tracking performance.

Datasets: We utilise two external datasets and a custom-developed dataset with a simplified version of *Piano Phase*. The *Candombe* dataset has 35 full-length songs with variable durations that accumulate to almost 2.5 hours [11]. The *Bambuco* dataset features two sets of ground-truth annotations corresponding to the predominant meters [12]: 3/4 and 6/8, referred to as *Bambuco (simple)* and *Bambuco (compound)* respectively.

Fig. 4: Musical score of the simplified version of *Piano Phase*.

To address the significant challenges *Piano Phase* presents for human annotators attempting to accurately annotate the beat "by ear", we created a simplified version (as depicted in Fig. 4) of the composition using a PureData patch. This patch produced two streams of 12 MIDI notes played at slightly different *tempi*, and the audio was obtained using a piano synthesizer. Ground-truth beat annotations were generated for each stream, assuming a 6/8 time signature (thus adopting the dotted quarter note (♩.) as the beat, as inferred from the original score). The score of this simplified rendition is shown in Fig. 2. Our primary experimental objective is to assess the ability of our beat-tracker to synchronise with each of the *tempi* present in the music. To achieve this, the custom dataset is composed of two files, `pianophaseM_A` and `pianophaseM_B`, representing the mixed audio (`M:A+B`) and ground-truth annotations for streams `A` and `B`.

4 Results

Beat Tracking in Uruguayan *Candombe*: Fig. 5 provides a summary of the overall results. A clear improvement in accuracy scores is observed across all fine-tuning configurations when compared to the baseline (`bsl`). Exceptions arise with configurations exclusively utilising DBN-parameterisation techniques (`pt` and `tg`), which yield scores similar to the baseline. Quantitatively, the best-performing configuration (`ft`) elevates the mean F-measure score from 0.280 to 0.952 when excluding the fine-tuned region (`testRes`), and from 0.334 to 0.956 when considering the entire file extent (`fullRes`).

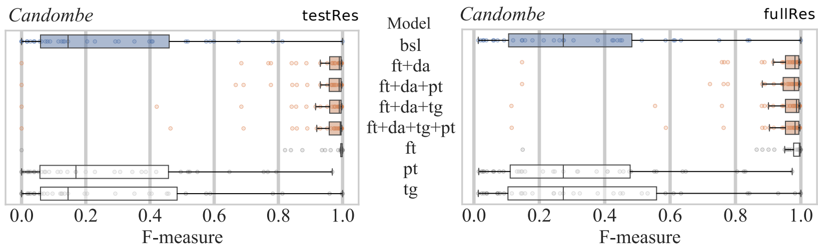

Fig. 5: Distribution of F-measure scores by configuration for the *Candombe* dataset.

When examining the annotation-correction workflow detailed in Table 1, it is observed that the Annotation Gain (Ag) improvements are marginally less than those of the F-measure. This indicates that the shift operation plays a minor role in this dataset's annotation workflow. However, the results demonstrate that our method significantly enhances efficiency. The number of operations (#ops) required to correct beat detections drops from 12,912 in the baseline to just 1,904 with the `ft` configuration. Given that there are 19,136 total beat annotations in the *Candombe* dataset, this means that the `ft` configuration requires only about 10% of the total beats to be corrected, achieving a reduction of approximately 85% from the baseline. Even accounting for the required user annotations for fine-tuning (which amount to 4,757 in the current experimental scenario), the results demonstrate a compelling decrease in manual annotation effort.

Table 1: Mean of the Ae score and sum of the #det, #ins, #del, #shf, and #ops scores across the *Candombe* dataset for the main configurations. (`fullRes`)

Dataset	Model	Ae	#det	#ins	#del	#shf	#ops
Candombe	bsl	0.319	6,316	2,901	92	9,919	12,912
	ft+da	0.919	16,688	1,885	181	*561*	2,632
	ft+da+pt	0.915	16,575	1,997	*178*	563	2,739
	ft+da+tg	0.922	16,892	*1,504*	190	739	2,437
	ft+da+tg+pt	**0.923**	**16,903**	*1,504*	188	726	*2,421*

Beat Tracking in Colombian *Bambuco*: As summarised in Fig. 6, all primary fine-tuning configurations outperform the baseline (`bsl`). Results are consistent across both settings (`testRes` and `fullRes`), revealing notable F-measure improvements: around 25 percentage points (p.p.) for the simple metre and close to 30 p.p. for the compound metre datasets. The `ft+da+tg` configuration emerges as the standout performer in both scenarios. Although each of the techniques (`ft`, `pt`, and `tg`) yields different contributions individually, their combined implementation is what truly augments performance.

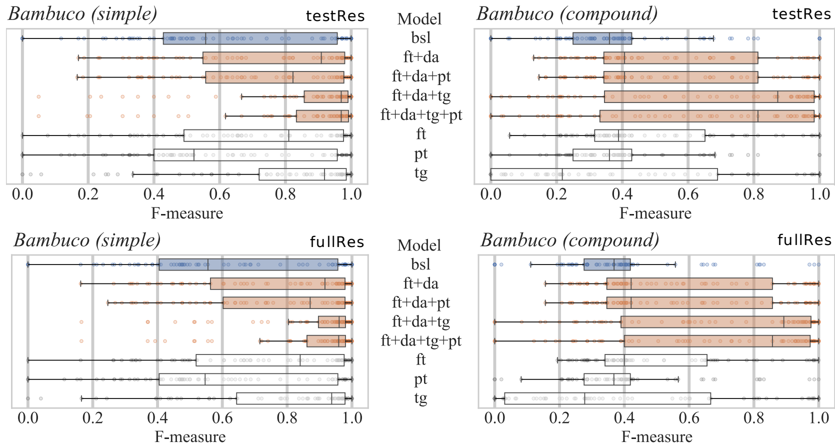

Fig. 6: Distribution of F-measure scores by configuration for the *Bambuco* datasets.

Table 2 shows Ae gains slightly outpacing F-measure, illustrating a greater relevance of the shift operation in this setting. Compared to the baseline, the `ft+da+tg` setup in simple metre trims beat estimate correction operations (#ops) by two-thirds (455 *vs* 1,610). For the compound subset, correct detections (#det) almost double in the optimal setting (from 899 to 1,665), underscoring our method's enhancement over the state of the art.

Table 2: Mean of the Ae score and sum of the #det, #ins, #del, #shf, and #ops scores across the *Bambuco* datasets for the main configurations. (`fullRes`)

Dataset	Model	Ae	#det	#ins	#del	#shf	#ops
Bambuco (simple)	bsl	0.556	1,756	1,110	*60*	440	1,610
	ft+da	0.726	2,439	602	91	*265*	957
	ft+da+pt	0.718	2,428	588	94	291	972
	ft+da+tg	**0.869**	**2,990**	*12*	138	306	*455*
	ft+da+tg+pt	0.866	2,978	*12*	137	316	465
Bambuco (compound)	bsl	0.338	899	424	410	947	1,781
	ft+da	0.509	1,319	285	340	665	1,292
	ft+da+pt	0.513	1,322	285	332	663	1,282
	ft+da+tg	**0.685**	**1,665**	*63*	62	*541*	*667*
	ft+da+tg+pt	0.671	1,640	73	61	557	691

Beat Tracking in Steve Reich *Piano Phase*: The primary experimental objective is to evaluate the capability of our beat-tracking method in synchronising with distinct *tempi* present in this musical piece. When referencing stream *A* or *B*, we are essentially assessing the beat tracker's ability to tune into each stream's tempo. This task, which is already immensely challenging for most humans, i.e., allowing themselves to align with one tempo while ignoring conflicting ones, presents an even more formidable test for an automatic beat tracker. Given this complexity, any advancement in performance, even if slight, can be considered significant. With this perspective, we now delve into the results obtained from our experiments.

From Fig. 7, results indicate improvements across all fine-tuning configurations when compared to the baseline for both streams (*A* and *B*). The F-measure score rises from approximately 0.2 to 0.7 across the main configurations. The role of fine-tuning (`ft`) is prominent, emerging as a key factor in performance enhancement. However, a more constrained adaptation to stream *B* is also apparent, an aspect we currently lack comprehensive data to fully elucidate. Another aspect worth further investigation is the observed efficiency of the adaptive processor type (`pt`) over the tempo guide (`tg`). This observation is somewhat counterintuitive, given that the primary goal of this beat tracking method aims to synchronise with conflicting, yet stable, *tempi*.

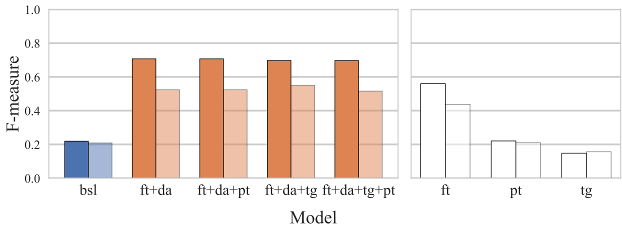

Fig. 7: F-measure *vs* model for *Piano Phase* (*left*:`pianophaseM_A`; *right*:`pianophaseM_B`).

As represented in Fig. 8, a closer examination of specific musical segments for stream *A* is provided. This figure offers a comparative perspective between the baseline approach and our best-performing configuration (`ft+da`). The superiority of the fine-tuned configuration over the baseline is evident across most parts of the musical segment. Notably, the beat estimates are accurate up to nearly bar *6* (or up to 68 seconds to be precise). However, around bar *6*, signs of desynchronisation emerge, with the imprecise predictions persisting until bar *8*. In this specific range, the baseline method manages to hold a slight edge over our approach by correctly identifying certain beats. In terms of the annotation-correction workflow, we see that the state-of-the-art correctly estimates 40 beats, while our fine-tuned configuration improves this count significantly, estimating 105 correctly. Even considering the required 19 user annotations for the fine-tuning segment, this is a notable improvement with such challenging material.

However, it is important to place the results obtained in the appropriate context. When comparing our method with non-adaptive beat trackers, including the current state-of-the-art, we recognize that this is not an even comparison. Most traditional beat trackers are designed for music that adheres to a single tempo, and data-driven methods have not been exposed to similar training examples, as *polytempo* is absent from standard datasets. Despite these differences, it remains logical to use a baseline for performance assessment. Our focus is in demonstrating that with minimal user input, our approach can leverage the model's general knowledge and adapt to music with rhythmic dissonance. This showcases the versatility of our approach and its applicability in diverse musical scenarios.

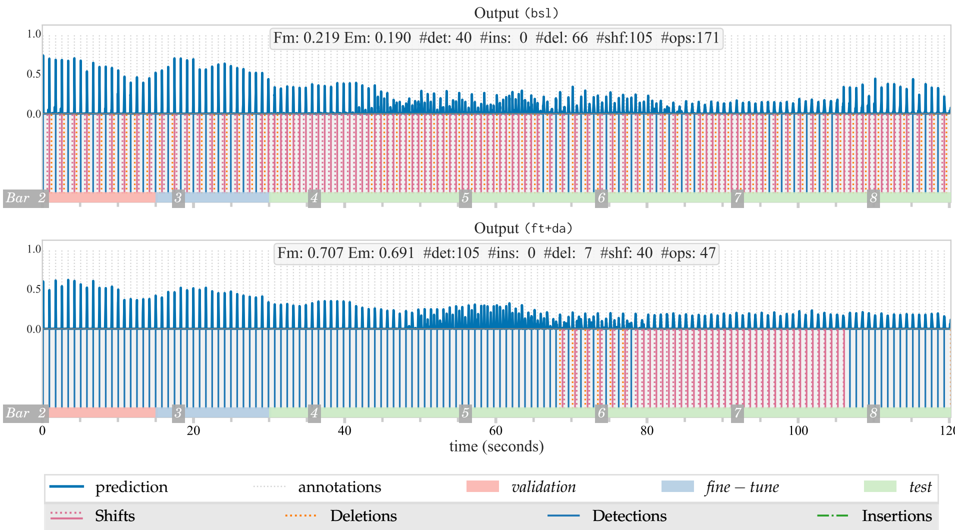

Fig. 8: Detailed analysis for *pianophaseM_A* (Mixed audio and annotations for stream *A* tempo).

5 Conclusions

In this study, we presented a user-centric approach to beat tracking designed specifically for challenging music signals. By leveraging concise user-annotated regions, our method significantly enhanced the performance of current state-of-the-art beat tracking, especially in environments dominated by complex rhythms. The rhythmic intricacies of Colombian *Bambuco*, Uruguayan *Candombe*, and Steve Reich's *Piano Phase* were put under scrutiny. These music forms represent, in order, the phenomena of *polyrhythm*, *polymetre*, and *polytempo*.

Among the notable results, for *Candombe*, our approach achieved an excess of 3-fold improvement over existing techniques. In the case of *Bambuco*, the performance was enhanced by approximately 25 p.p. for the simple metre and neared 30 p.p. for the compound metre datasets. With Reich's *Piano Phase*, even though the F-measure score escalated 50 p.p., our primary objective was to underscore our method's capability in handling the extreme challenges posed by *polytempo*. To the best of our knowledge, this study is the first to attempt beat tracking of a musical composition with such compositional technique.

While these results are promising, it is essential to interpret accuracy variations carefully and circumscribe them to the scope of our investigation. Looking forward to future research directions, the exploration of extended musical segments, enriched with a diverse set of fine-tuning parameters, could provide more profound insights into *polytempo* adaptability. Though this study's scope was restricted, it introduces promising methodologies for situations where traditional techniques might not be as effective.

In summary, our research demonstrated the potential of transfer learning and user-driven adaptation for beat tracking in rhythmically complex musical contexts. Using minimal user feedback, we enhanced the state-of-the-art model, enabling its adaptability to challenging musical scenarios and underscoring its utility for specific applications, notably musicological analysis. Our research reach extends past beat tracking, touching upon rhythm-focused tasks such as metre determination and downbeat tracking. Yet, our user-centred approach suggests even wider application across various MIR tasks, beyond computational rhythm analysis. Given the inherent ambiguity in music signals, integrating a user-centric viewpoint is pivotal in integrating subjectivity and accurate analysis.

While our findings represent an encouraging step forward, there remains much to explore in this domain. We hope this study serves as a starting point for future endeavours, aiming to refine adaptive strategies and the human-in-the-loop paradigm. Ultimately, our goal is to promote the development of MIR tools capable of effectively handling a wider range of musical traditions, fostering inclusivity and a deeper appreciation of the world's rich musical heritage.

References

1. K. Agawu and V. K. Agawu. *African Rhythm: A Northern Ewe Perspective*. Cambridge University Press, 1995.
2. S. Böck and M. E. P. Davies. Deconstruct, Analyse, Reconstruct: How To Improve Tempo, Beat, and Downbeat Estimation. In *Proc. of the ISMIR*, pages 574–582, 2020.
3. E. Cano, F. Mora-Ángel, G. A. L. Gil, J. R. Zapata, A. Escamilla, J. F. Alzate, and M. Betancur. Sesquialtera in the Colombian Bambuco: Perception and Estimation of Beat and Meter – Extended version. *Trans. of the ISMIR*, 4(1):248–262, 2021.
4. O. Cornelis, M. Lesaffre, D. Moelants, and M. Leman. Access to ethnic music: Advances and perspectives in content-based music information retrieval. *Signal Processing*, 90(4):1008–1031, 2010.
5. M. E. P. Davies and S. Böck. Temporal convolutional networks for musical audio beat tracking. In *Proc. of the 27th European Signal Processing Conference*, 2019.
6. D. Fiocchi, M. Buccoli, M. Zanoni, F. Antonacci, and A. Sarti. Beat Tracking using Recurrent Neural Network: A Transfer Learning Approach. In *26th European Signal Processing Conf.(EUSIPCO)*, pages 1915–1919. IEEE, 2018.
7. M. Fuentes, L. S. Maia, M. Rocamora, L. W. Biscainho, H. C. Crayencour, S. Essid, and J. P. Bello. Tracking beats and microtiming in Afro-latin American music using conditional random fields and deep learning. In *Proc. of the ISMIR*, pages 251–258, 2019.
8. E. Gómez, P. Herrera, and F. Gómez-Martin. Computational Ethnomusicology: perspectives and challenges. *JNMR*, 42(2):111–112, 2013.

9. F. Gouyon and S. Dixon. A Review of Automatic Rhythm Description Systems. *Computer Music Journal*, 29(1):34–54, 2005.
10. A. Holzapfel, M. E. P. Davies, J. R. Zapata, J. L. Oliveira, and F. Gouyon. Selective sampling for beat tracking evaluation. *IEEE Trans. on Audio, Speech and Language Processing*, 20(9):2539–2548, 2012.
11. L. Jure and M. Rocamora. Microtiming in the rhythmic structure of Candombe drumming patterns. In *Fourth Int. Conf.on Analytical Approaches to World Music (AAWM 2016)*, pages 1–5, 2016.
12. F. Mora-Ángel, G. L. Gil, E. Cano, and S. Grollmisch. ACMUS-MIR: A new annotated data set of Andean Colombian music. In *7th Int. Conf.on Digital Libraries for Musicology (DLfM)*, 2019.
13. L. Nunes, M. Rocamora, L. Jure, and L. W. Biscainho. Beat and downbeat tracking based on rhythmic patterns applied to the Uruguayan candombe drumming. In *Proc. of the ISMIR*, pages 264–270, 2015.
14. A. S. Pinto, S. Böck, J. S. Cardoso, and M. E. P. Davies. User-Driven Fine-Tuning for Beat Tracking. *Electronics*, 10(13):1518, 2021.
15. A. S. Pinto and M. E. Davies. Tapping Along to the Difficult Ones: Leveraging User-Input for Beat Tracking in Highly Expressive Musical Content. In R. Kronland-Martinet, S. Ystad, and M. Aramaki, editors, *Perception, Representations, Image, Sound, Music - 14th Int. Symposium, CMMR 2019, Marseille, France, Oct. 14-18, 2019, Revised Selected Papers*, volume 12631 of *LNCS*, pages 75–90. Springer, 2021.
16. A. S. Pinto and M. E. P. Davies. Towards user-informed beat tracking of musical audio. In *14th Int. Symposium on Computer Music Multidisciplinary Research (CMMR)*, pages 577–588, 2019.
17. A. S. Pinto, I. Domingues, and M. E. P. Davies. Shift If You Can: Counting and Visualising Correction Operations for Beat Tracking Evaluation. In *Extended Abstracts for the Late-Breaking Demo Session of the Int. ISMIR Conf.*, 2020.
18. J. M. Schechter, D. E. Sheehy, and R. R. Smith. Latin America. *Ethnomusicology*, 29(2):317, 1985.
19. A. Srinivasamurthy, A. Holzapfel, and X. Serra. Informed automatic meter analysis of music recordings. In *Proc. of the ISMIR*, pages 679–685, 2017.
20. G. Tzanetakis, A. Kapur, W. Schloss, and M. Wright. Computational Ethnomusicology. *J. of Interdisciplinary Music Studies*, 1(2):1–24, 2007.