Queen Mary University of London

Digital Music Research Network

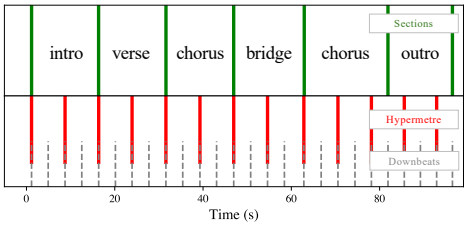
# Can downbeat trackers predict hypermetre?

Jose Alejandro Esquivel and Jordan B. L. Smith

j.esquiveldejesus@qmul.ac.uk

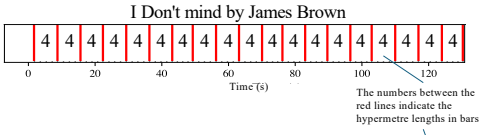
## Introduction

**Hypermetre** is the metric structure between the timescales of downbeats and sections in a piece of music.



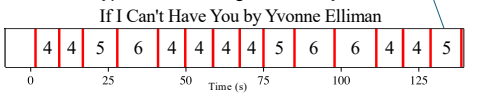

Hypermetre tracking systems face the same challenges as beat and downbeat tracking systems: changes in tempo.

There are songs with constant hypermetre (possibly most of them):




The numbers between the red lines indicate the hypermetre lengths in bars

However, hypermetre changes are very common:

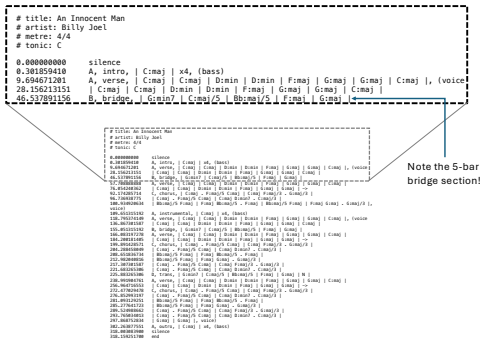



We want to develop custom algorithms to perform hypermetre tracking. First, though, we check whether existing tools can be adapted to perform the task.

In this work, we propose several methods to adapt downbeat tracking methods to predict hypermetre and evaluate them on the McGill Billboard dataset.

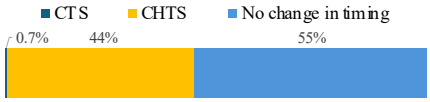
## Data

We are not aware of any datasets of hypermetre annotations. However, in the McGill Billboard dataset of chord labels, the line breaks in the annotation files appear to indicate the hypermetre.




In this dataset, 0.7% of 890 songs exhibit a change in time signature (CTS).

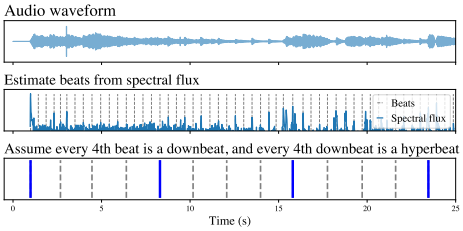
This contrasts with 44% of the songs that appear to have a change in hyper-time signature (CHTS).




## Methods

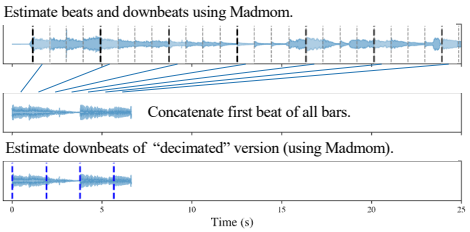
### The baseline method

This method serves as a simple baseline for performance floor. We assume 4/4 for the time signature and the **hyper-time-signature**.



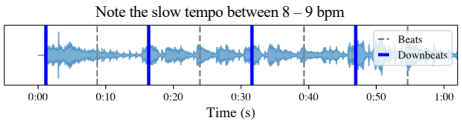

### The decimated method

This is a first attempt to use an existing beat tracker (Madmom) to perform hypermetre tracking.



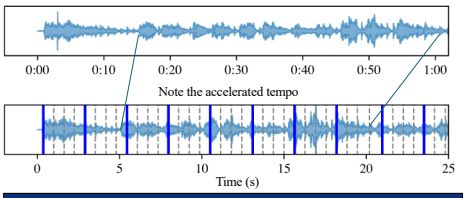

### The slow tempo method

In this method, we also use Madmom, but we set its parameters to expect a tempo between 3 bpm and 10 bpm.




### The accelerated method

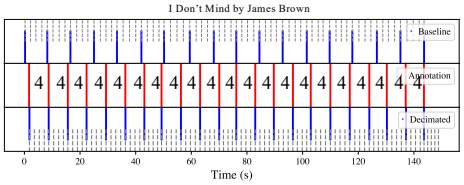
We also used Madmom for this method, but instead of setting the tempo, we accelerated the audio of the song (using librosa) by a factor of 3.




## Results

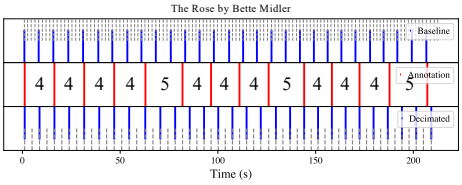
We applied all 4 methods to a 26-song subset (CTS =11%, CHTS=69%) of the McGill Billboard dataset and evaluated the predicted hyperbeat using precision, recall and F-measure with 0.5s tolerance.

"I Don't Mind" has a constant hyper-time-signature of 4 and shows perfect prediction with the decimated method (F=1.0), but for the other methods, the performance falls to F ≤ 0.2.

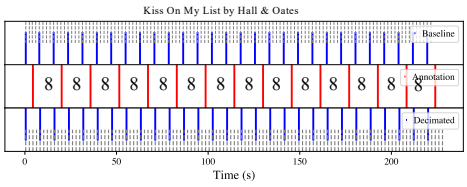



## Results (continuation)

In "The Rose", the baseline and decimated methods both produce predictions that are accurate at the start but that degrade when the hyper-time-signature changes (i.e., 4–5–4). For this song, the accelerated and the decimated methods both perform the same (F=0.37).

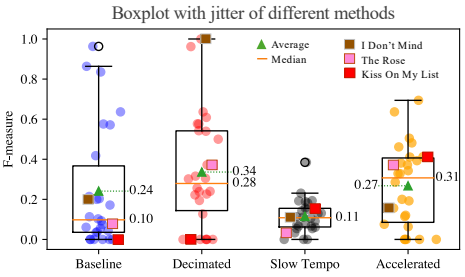



In "Kiss On My List", a song with constant hypermetre of 8, the decimated and baseline methods both fail (F=0), due to incorrect phase. The accelerated method performs the best for this song (F=0.41).




### Results overall

The decimated method performs the best on average (F=0.34), followed by the accelerated method (F=0.27). The baseline and the slow tempo perform alike (F ≈ 0.1). The decimated and baseline methods show larger spread of data compared to accelerated and slow tempo methods as shown in the next figure.




The factors affecting the performance of the methods are:

1. **Extra line breaks**: Some line breaks seem to be there to improve the legibility of the chord annotations rather than to indicate hyperbeats. The dataset needs careful validation.
2. **Changes in hyper-time-signature**: The hyper-time-signature exhibits changes with much more frequency than time signature changes, which the Madmom downbeat tracker cannot predict accurately.

## Conclusion

We found that Madmom, a system tuned to beat and downbeat information, can be used for hypermetre tracking with very limited quality.

To improve hypermetre tracking, we should account for music characteristics other than spectral flux (e.g., harmony and instrumentation).