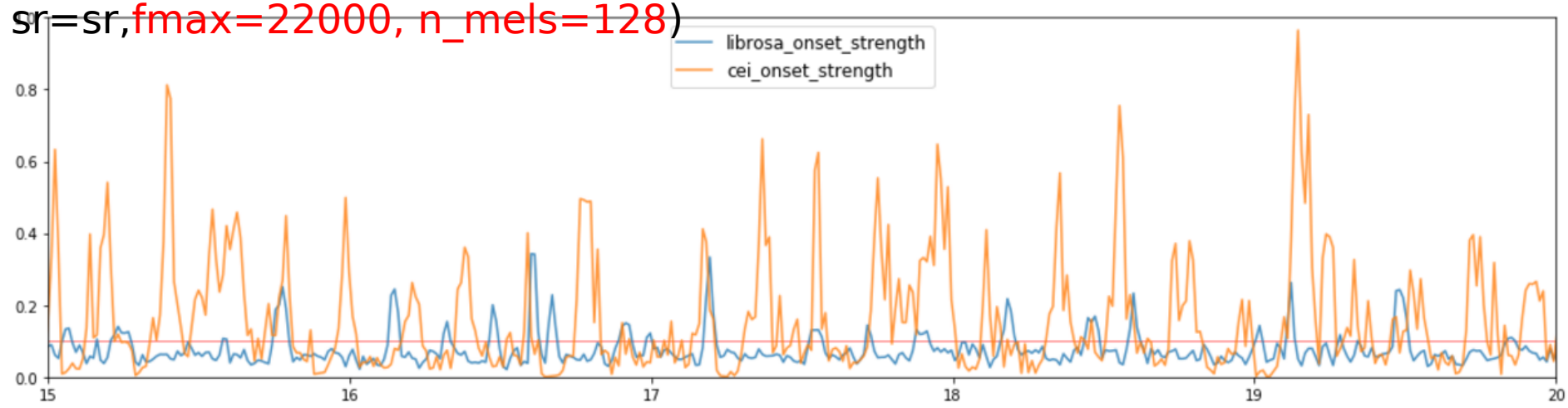
Data Analysis for Onset Detection

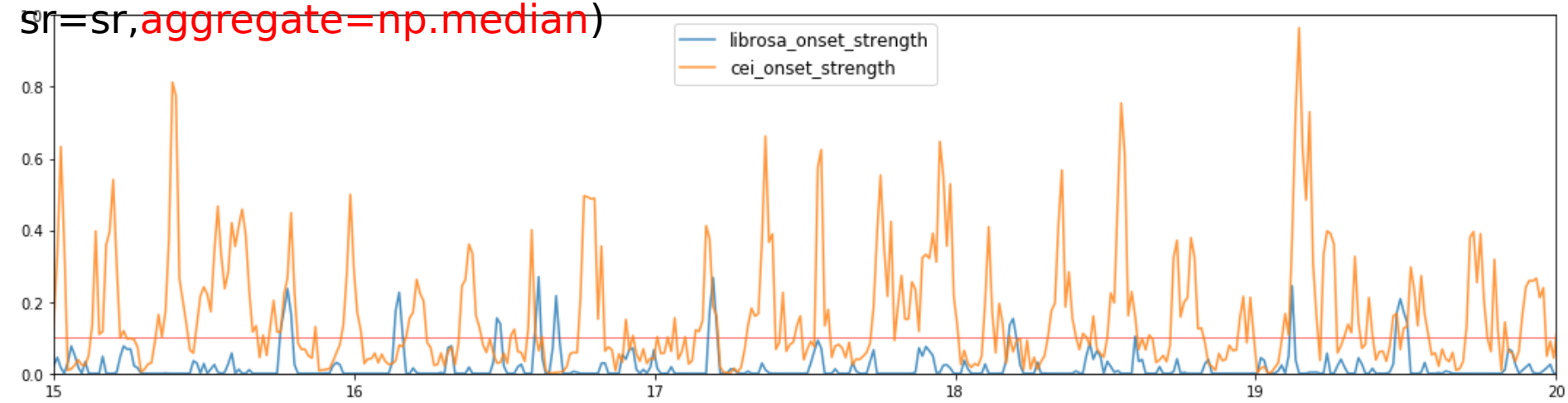
Alvin/Johnny

Onset strength output is different from librosa

onset_env = librosa.onset.onset_strength(y=audio_raw_data, sr=sr,fmax=22000, n_mels=128)



onset_env = librosa.onset.onset_strength(y=audio_raw_data, sr=sr,aggregate=np.median)

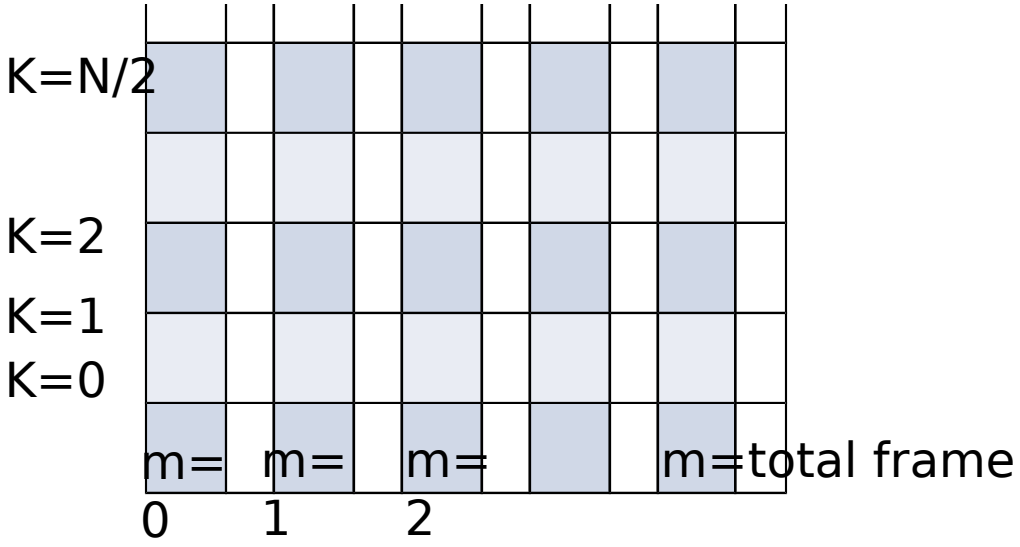


http://www.nyu.edu/classes/bello/MIR_files/3-novelty.pdf

- Use half-wave rectification to only take energy increases into account

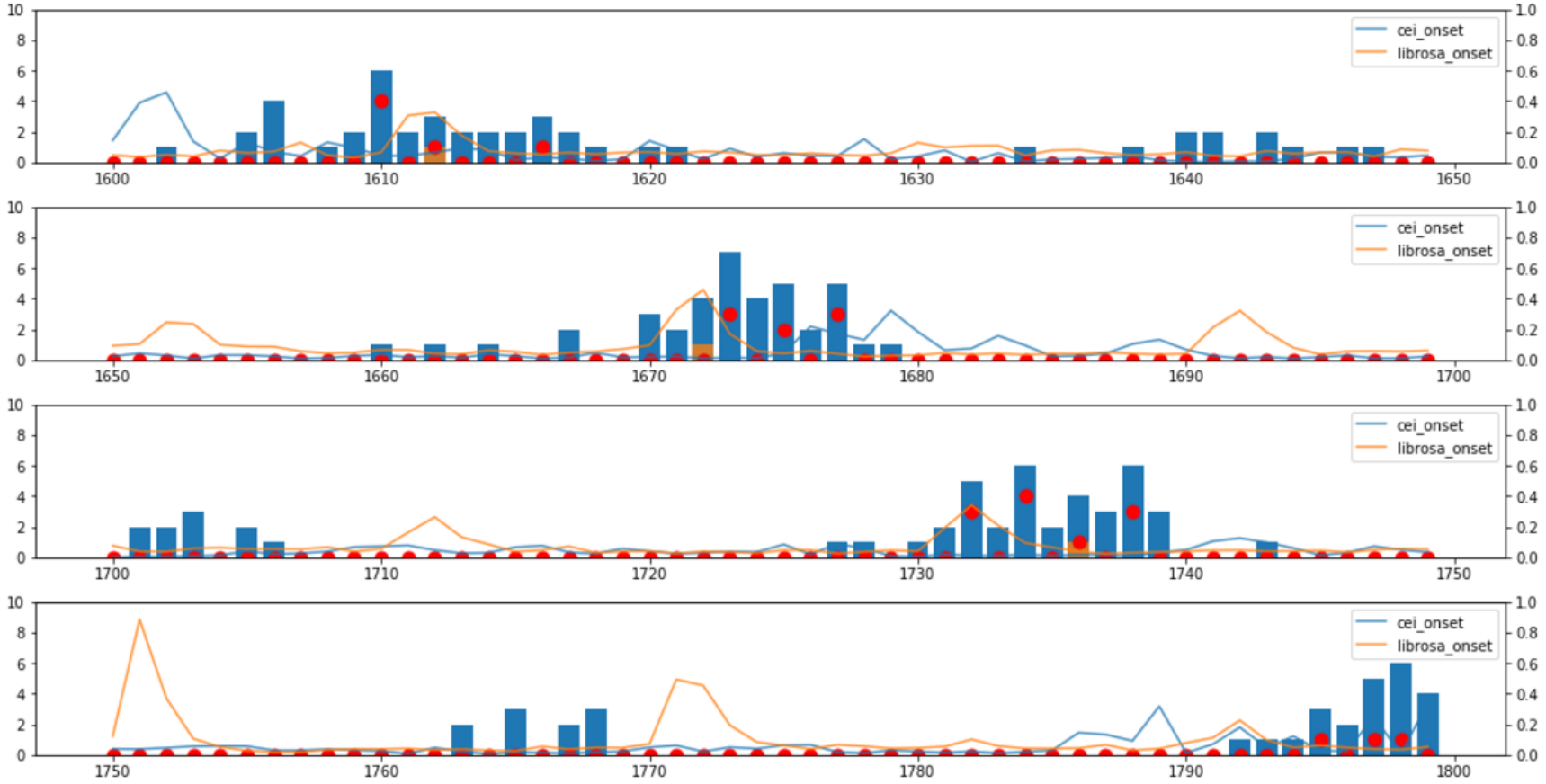
$$SF_R(m) = \frac{2}{N} \sum_{k=0}^{N/2} H\left(|X_k(m)| - |X_k(m-1)|\right)$$

$$H(x) = (x + |x|)/2$$




```
num_frame  = trainYA.shape[1]
trainX_sid =
trainX[sid*num_frame:sid*num_frame+num_frame,0:n_of_freq_bins]
trainXPD   = np.diff(trainX_sid,axis=0)
trainXPDH = np.maximum(np.zeros(trainXPD.shape),trainXPD)
```

Label Distribution
librosa onset_strength match better



- precision: 0.8132 - recall: 0.6653
- val_precision: 0.3823 - val_recall: 0.3017

Conclusion

- Librosa onset_strength is different from cei implementation
- Librosa onset_strength match user labeling output better
  - Basically, cei implementation is not wrong, maybe there is another tips i don’t know
- CEI LSTM model get the following result, maybe more data could improve P/R value
  - precision: 0.8132 - recall: 0.6653
  - val_precision: 0.3823 - val_recall: 0.3017
  - Trainging song 16, testing song 4.