Random tracking title

Till Raab, ..., ..., Jan Benda

September 24, 2018

Abstract

work in progress

Introduction

work in progress

Materials and Methods

Field site

The study site was a small stream called Rio Rubiano in the Colombian part of tropical grassland plain "Los Llanos" near San Martin, Province Meta. The recording site was an easy to access part of the Rio Rubiano near the Finca Altamira (3°76'52.70"N, 73°67'53.41"W) which also served as accommodation. The river bed consists of rocks with a diameter ranging from a few to up to 50 cm and the riverbank consists mainly of soil, rocks and the roots of the surrounding vegetation. The very location the recording equipment was installed was a part of the river where the river width is approximately 9 m and water depth is around 20 cm (distance between water surface and stone layer on the riverbed). The temperature of the clear water of Rio Rubiano fluctuated between 23 and 27 °C on a daily basis and showed a conductivity ranging from 2 μS/cm to 7 μS/cm. Data acquisition started in April 2016, i.e. during the start of the rainy season.

Field monitiring system

The recording system used to obtain our date is similar to the one used by [Henninger et al. (2018)] in the Republic of Panamá. It consists of a custom-build 64-channel electrode and amplifier system (npi electronics GmbH, Tamm, Germany) powered by 12 V car batteries. Signals detected by the electrodes (low-noise headstages embeded in epoxy resin (1 $\times$ gain, 10 $\times$ 5 $\times$ 5 mm)) were amplified by the main amplifier (100 $\times$ gain, 1st order high-pass filter 100 Hz, low-pass 10 kHz) before being digitalized with 20 kHz per channel with 16-bit amplitude resolution using a custom build computer with two digital-analog converter cards (PCI-6259, National Instruments, Austin, Texas, USA). Data acquisition and storage for offline analysis were managed by a custom software written in C++ (https://github.com/bendalab/fishgrid). The maximum of 64 electrodes mounted on 8 PVC tubes were arranged in an 8 by 8 electrode grid (50 cm spacing) covering an area of 350$\times$350 cm. All 64 electrodes were used throughout the whole recording period. Each PVC tube, equipped with 8 electrodes got tied to a rope crossing the river, forming a structure allowing small shifts in electrode distance but being resilient to destruction by rapidly changing environmental factors, i.e. rising water levels after heavy rainfall.

Extraction on EOD frequencies

Tracking of individual EODs

In order to track individual EOD frequency traces for each individual recorded we developed an algorithm based on Python3 using two independent signal variables to reliable assign EOD frequency traces.

Δ-*EODf (Electric organ discharge frequency difference)*

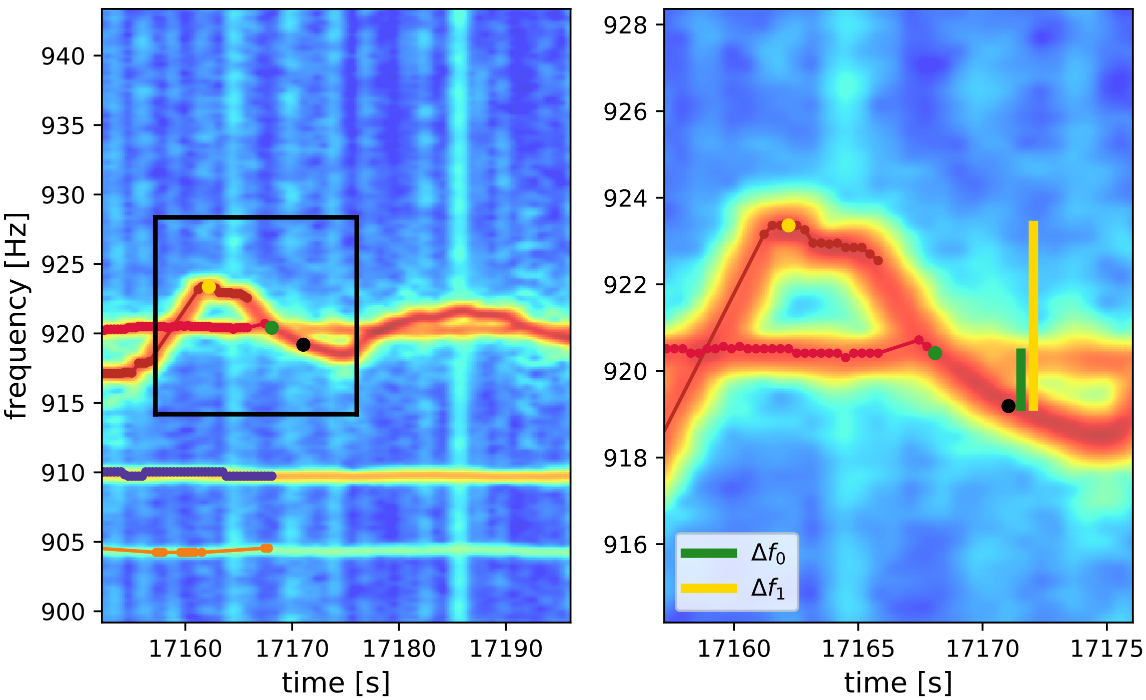

Figure 1: Random caption

Since, the EOD frequency of wave-type weakly electric fish represent on of the most stable oscillating signals known across natural systems and, thus, keeps stable for long periods of time it seems unambiguous to use this signal parameter as main tracking criterion. However, tracking individual EOD frequencies over long periods of time gives rise to further challenges. When wave-type weakly electric fish produce communication signals, such as EOD frequency rises, EOD frequency traces of different individuals frequently cross each other, sometimes multiple times within a short time window. This specially occurs in larger groups of wave-type weakly electric fish. These EOD frequency trace crossings give rise to several algorithmic problems. First, detecting both peaks during the Powerspectrum analysis in the very moment of the EOD frequency traces crossing often fails resulting in missing datapoints for one EOD frequency trace. As a result, correct tracking of EOD frequency traces, only based on EOD frequency comparison, is at chance level during these crossing events.

Δ-*F (Field difference)*

To address EOD frequency tracking errors arising from crossing EOD frequency traces, e.g. during the events of EOD frequency rises, we use the individual absolute fields properties as second tracking parameter. Due to our multi electrode recoding setup we are able to estimate the strength of each individual EOD signal at multiple locations within our electrode-grid by extracting the power of the according EOD frequency in the Powerspectrum of each

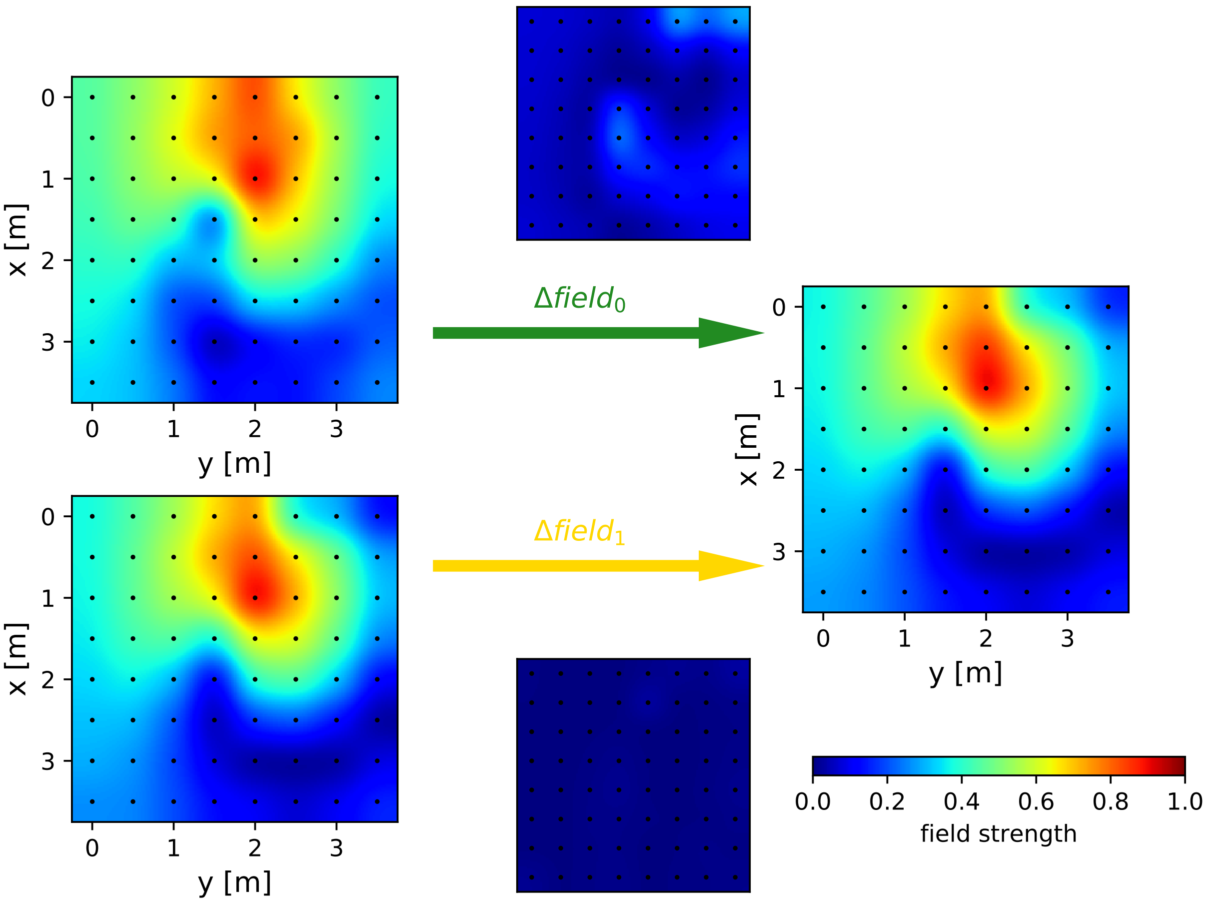

Figure 2: Random caption

electrode. These two-dimensional representations of the electric fields vary between individuals depending on their very location within the electrode-grid. After normalizing the individual electric fields to eliminate the impact of absolute field strength the obtained field proportions can be used as a second tracking parameter by calculation of the difference between two field proportions using the mean-square-error of different field proportions.

Error values composed from Δ*-EODf and* Δ*-Field*

The simple comparison of EOD frequency difference and field structure difference is not sufficient to determine the likelihood of two signals originating from the same individual. Therefore, relative EOD frequency errors and relative field errors, both ranging between 0 and 1, are calculated.

Frequency error determination

With respect to EOD frequency differences we assume EOD frequency differences of above 1 Hz to be equally unlike to originate from the same individual and, thus, result in the maximum relative EOD frequency error. EOD frequencies below 1 Hz result in smaller relative EOD frequency errors and are calculated from a Boltzmann-function resulting in smaller relative EOD frequency difference the lower the real EOD frequency difference is.

Field error determination

Considering the difference in field structure the absolute field structure errors are highly dependent on the amount of electrodes used in the recording setup. Therefore, to estimate the relative field structure error we first estimate

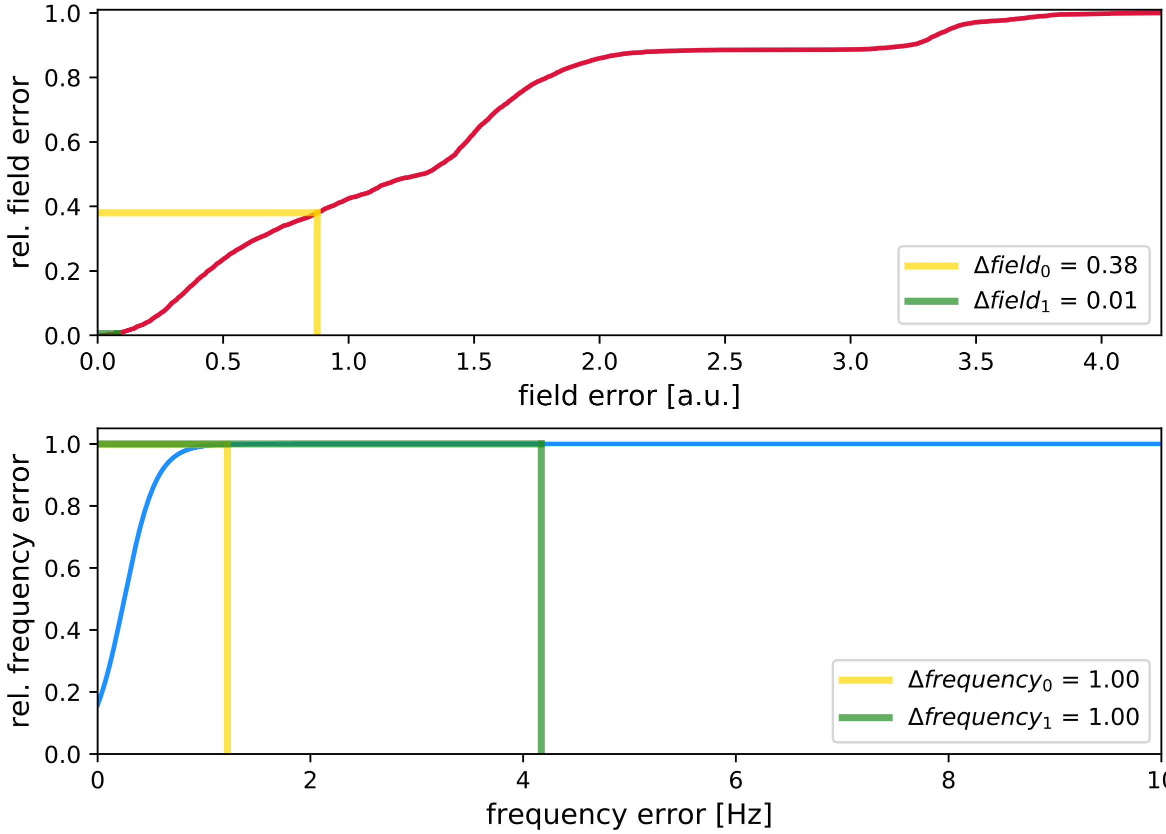

Figure 3: Random caption

the distribution of possible field structure errors in a 30 seconds window around the currently datapoints of interest. These possible field structure errors are define as those field structures with a smaller EOD frequency difference than 10 Hz. Deducted from the distribution of possible field structure differences the relative field structure difference of two field structures is the proportion of smaller field structure differences in the distribution of possible field structure differences.

Total error definition

The absolute error between two signals is calculated using a cost-function evaluating both, relative EOD frequency error and relative field structure error. Since frequency changes of several Hz within EOD frequency traces are possible due to the uttering of communication signals like EOD frequency rises and rapid spatial changes comparably uncommon (see Results) we use the cost-function displayed in equation 1 to estimate the total error value between different detected EOD signals.

Assign temporal EOD frequency traces

To enable the analysis of recordings with limitless duration, the actual tracking algorithm is two-staged. First, we assign so called temporal identities for EOD signals detected in a 30 seconds window. Therefore, we calculated the total error for every possible connection within this 30 seconds window. Besides the limitation of a maximum EOD frequency difference of 10 Hz the possible EOD signal pairs are limited by a maximum compare range of 10 seconds, i.e. two signals that shall be connected show a maximum time difference of 10 seconds. According to the obtained error values temporal identities are assembled starting from the smallest total error representing the best connection to the largest total error representing the least good connection. Connections that would interfere with already existing temporal identities are not made since already made connections are based on smaller total errors and therefore are more likely to be correct. The resulting temporal identities are based on the best possible connections, but only the centered 10 seconds of the 30 seconds window represent valid connections since the connections within the head and tail 10 seconds did not take into account all possible connections within $\pm$ compare range.

Running connection

The assembly of the center 10 seconds of the temporal identities, containing valid connections, and already tracked real identities, again is based on total errors between respective EOD signals. Total error values between signals of already assigned read identities and signals within the centered 10 seconds of temporal identities are identified and, again, assembled based on the total error values preferring lower total errors before larger ones. Temporal identities, which could not be assigned to a already existing real identity form a new real identity. The window to identify temporal identities is shifted by the compare range, i.e. 10 seconds, and the identification of temporal identities and their assignment to already tracked real identities continuous until the end of the recording is reached.

Results

d-EODf and d-Field of same identity signals vs. non-same identity signals

Roc analysis

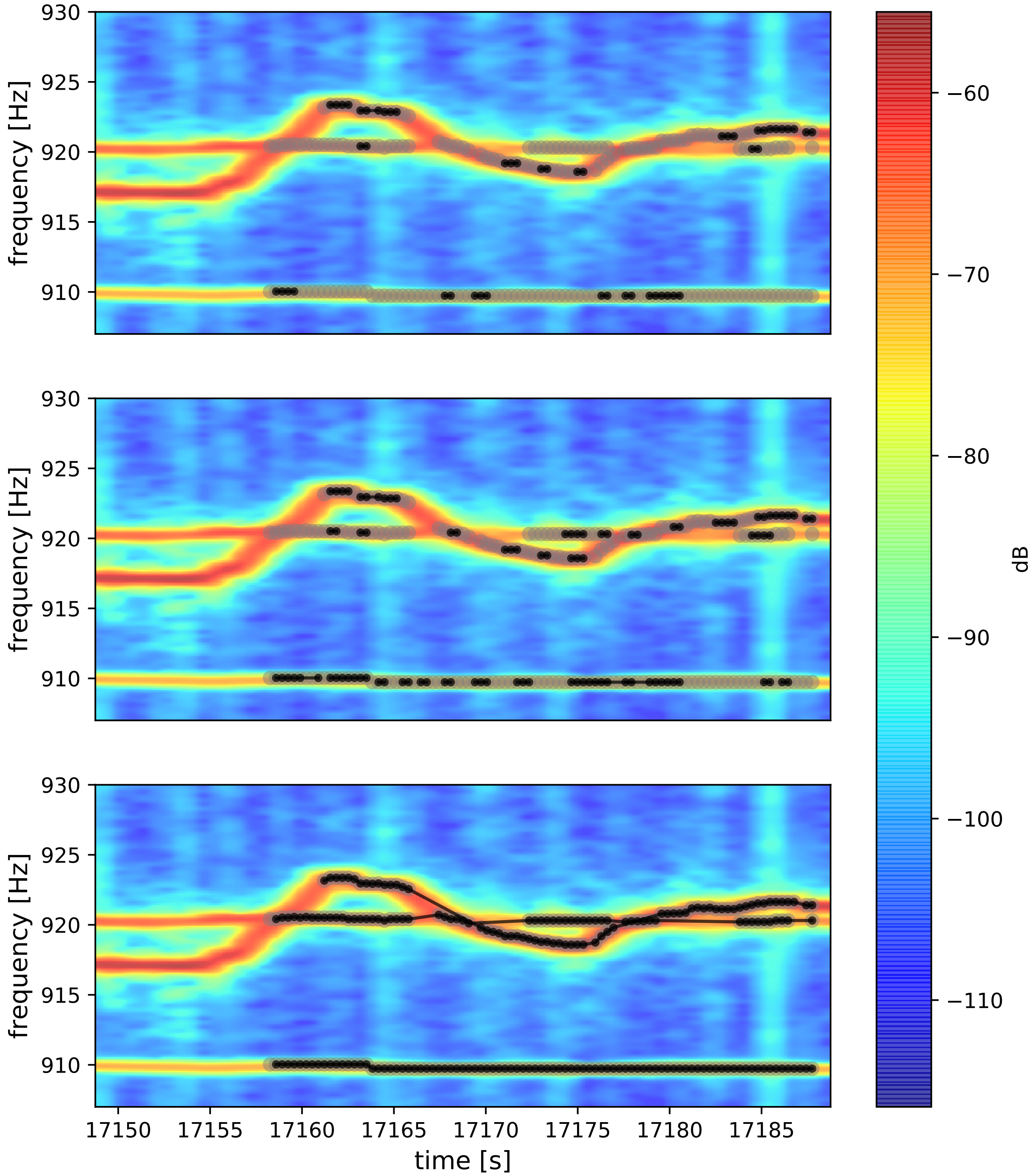

Figure 4: Random caption

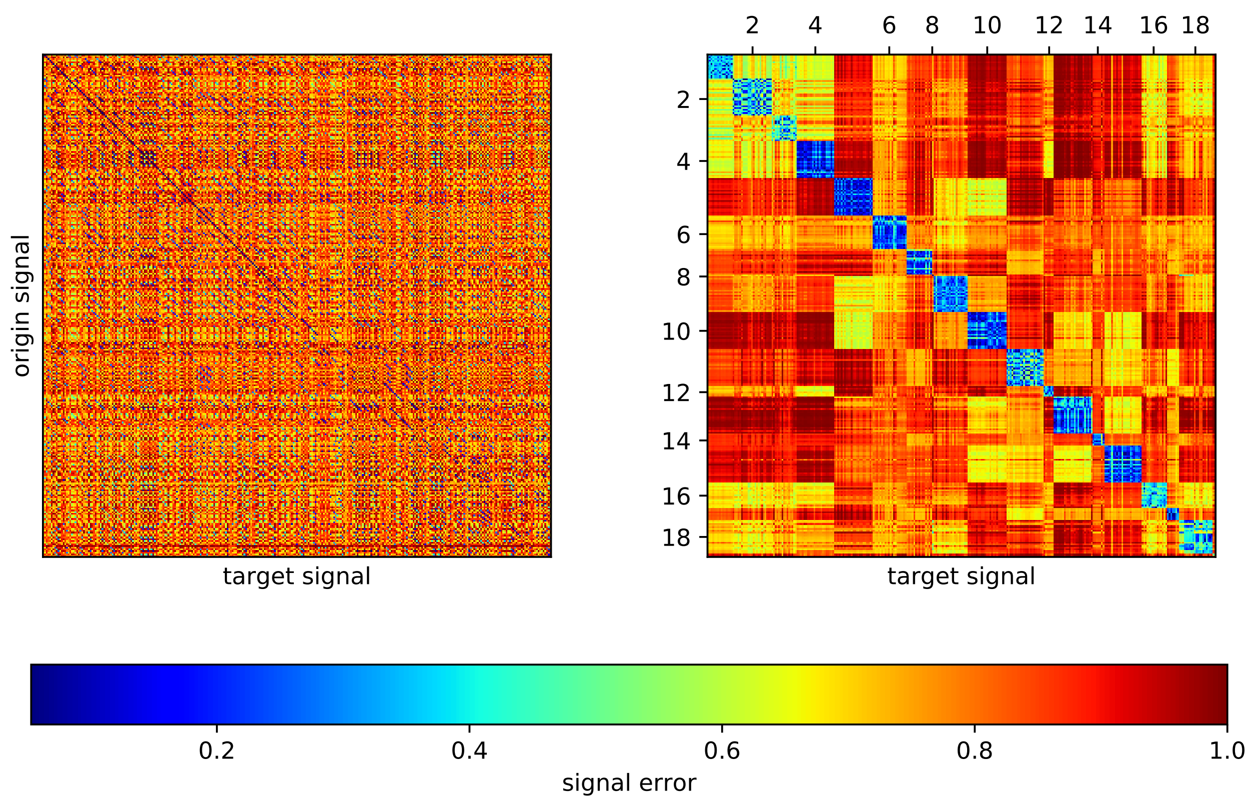

Figure 5: Random caption

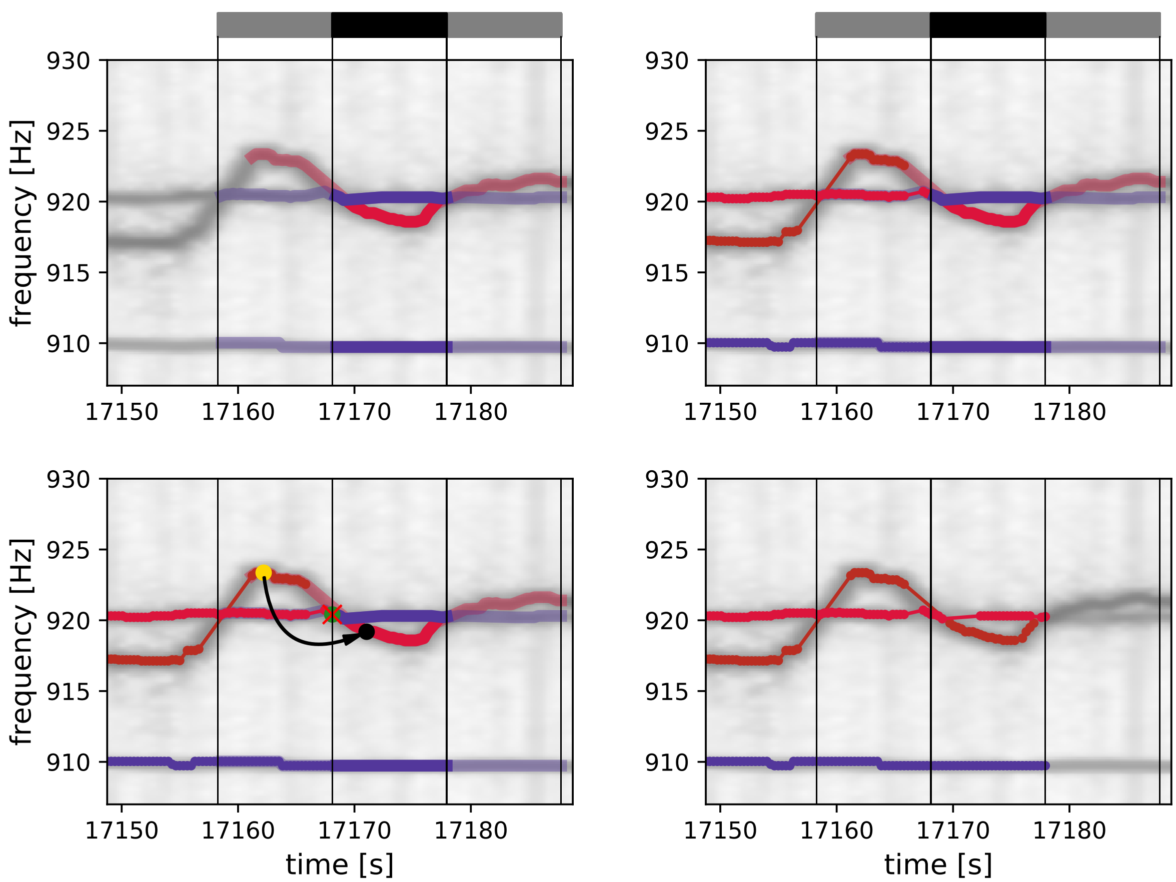

Figure 6: Random caption

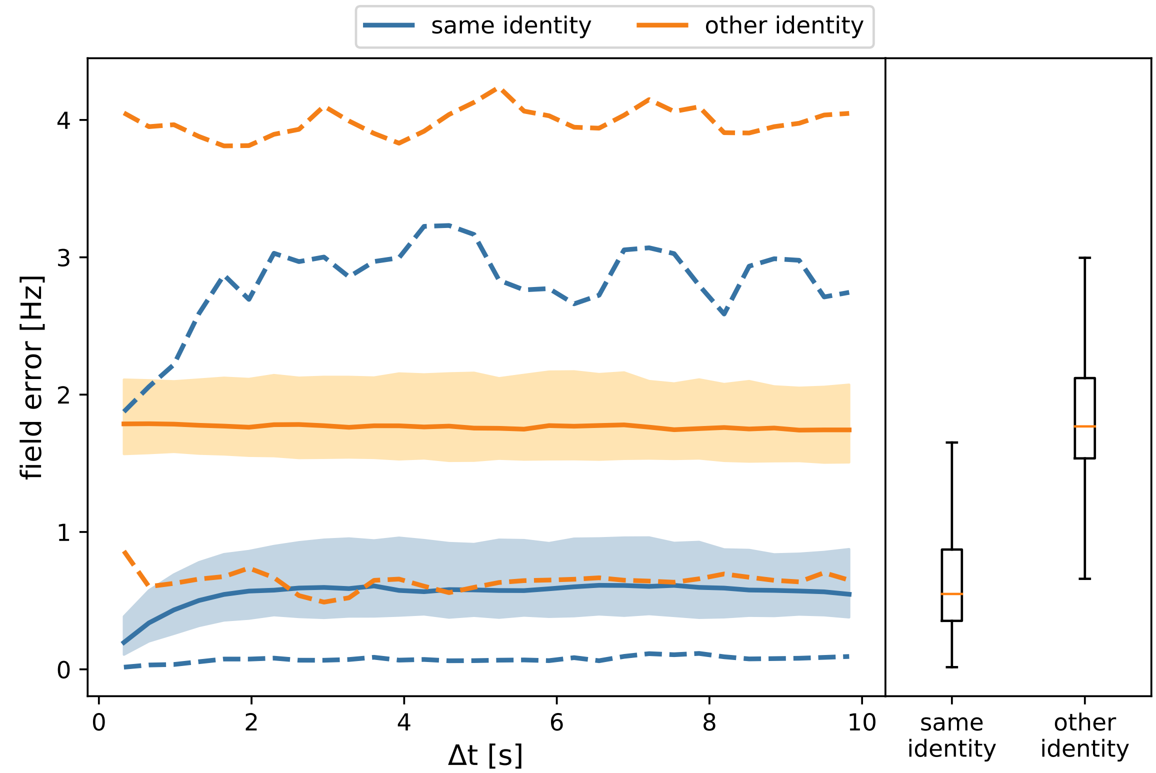

Figure 7: Random caption

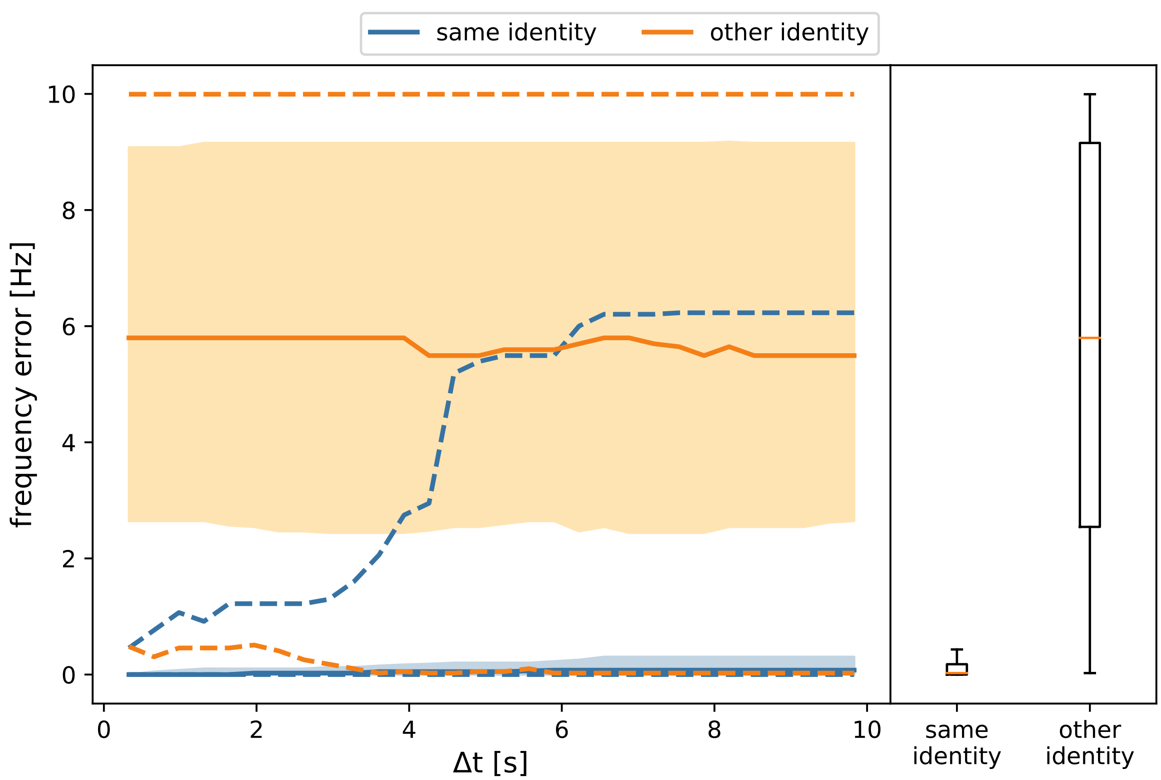

Figure 8: Random caption

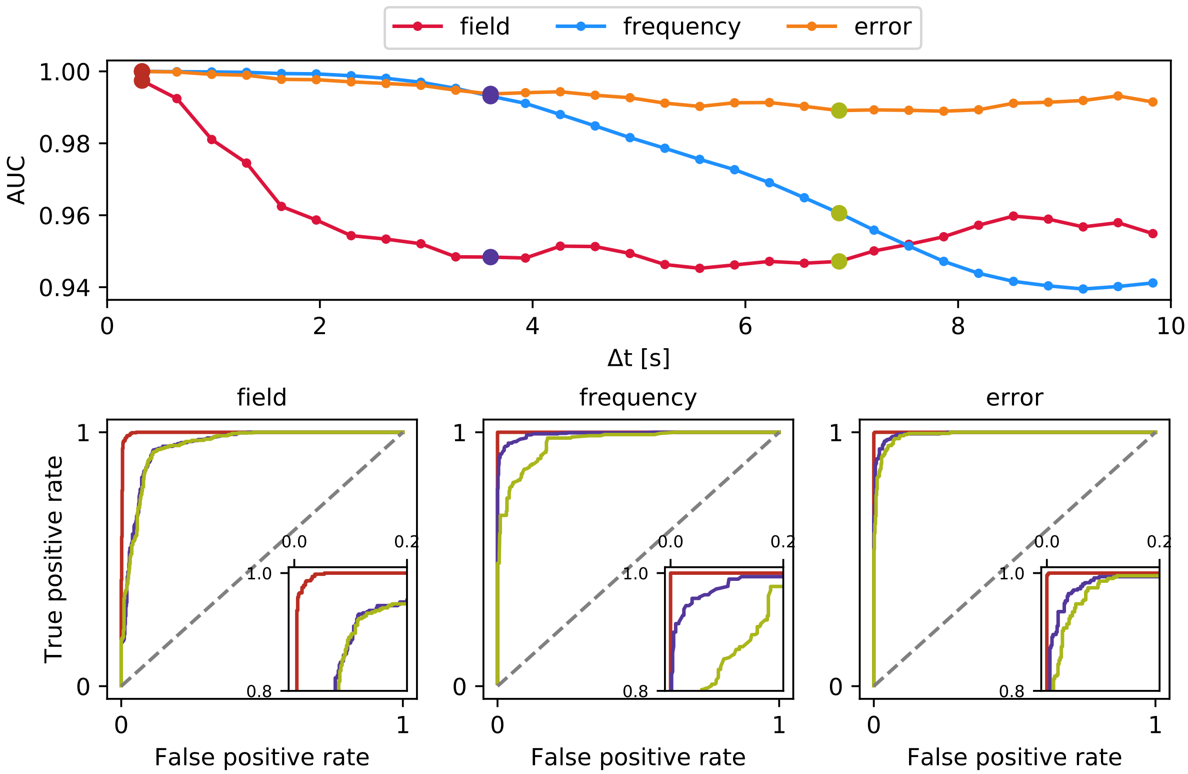

Figure 9: Random caption