



Bird Species Classification

The International Challenge from Amazon Rainforest

LSIS SABIOD & Life Clef Lab.

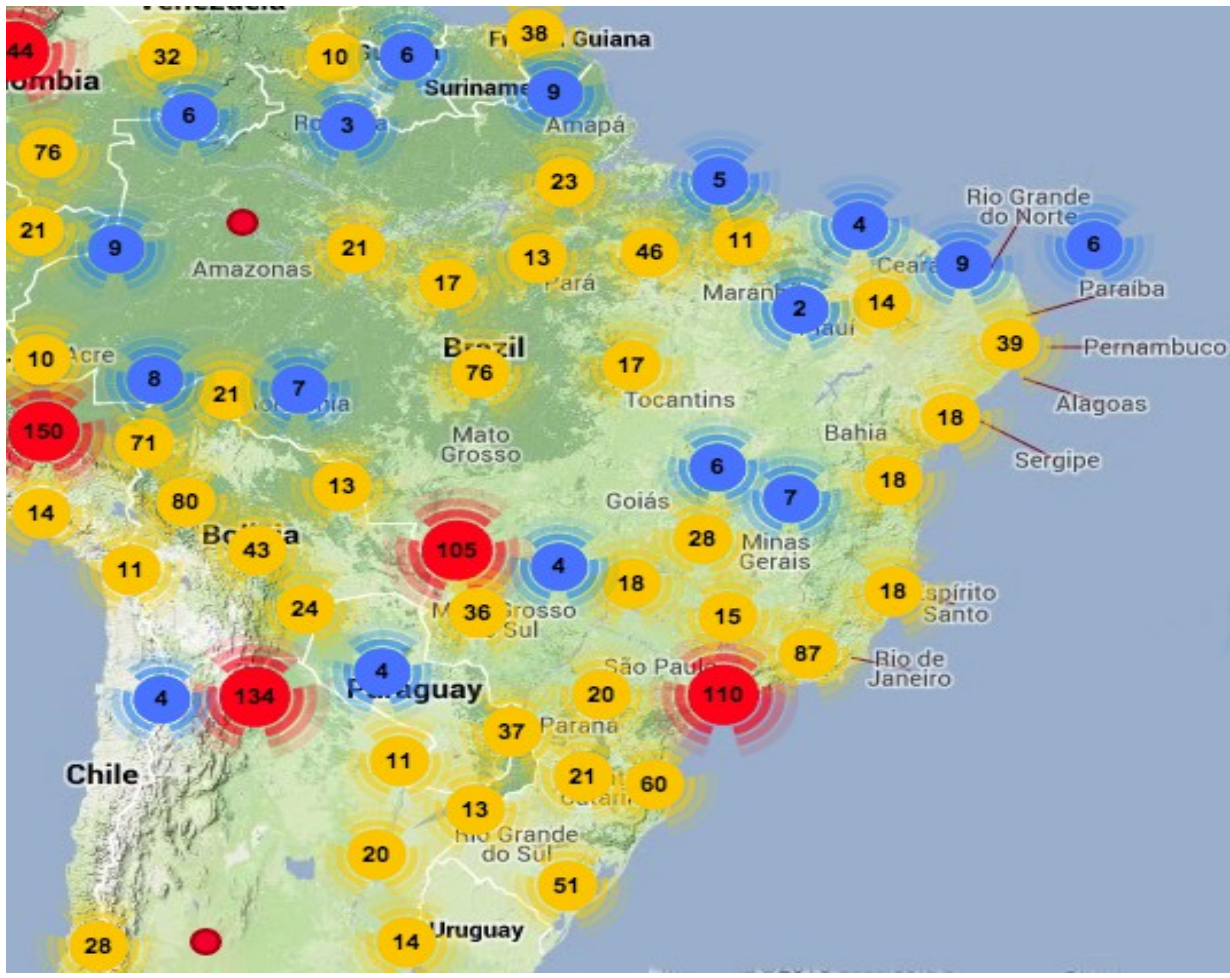
H. Glotin (Toulon), A. Joly (INRIA), Xeno Canto.

<http://glotin.univ-tln.fr>

University of Toulon, CNRS LSIS UMR 7296 - Institut universitaire de France



XENO CANTO citizen science dataset

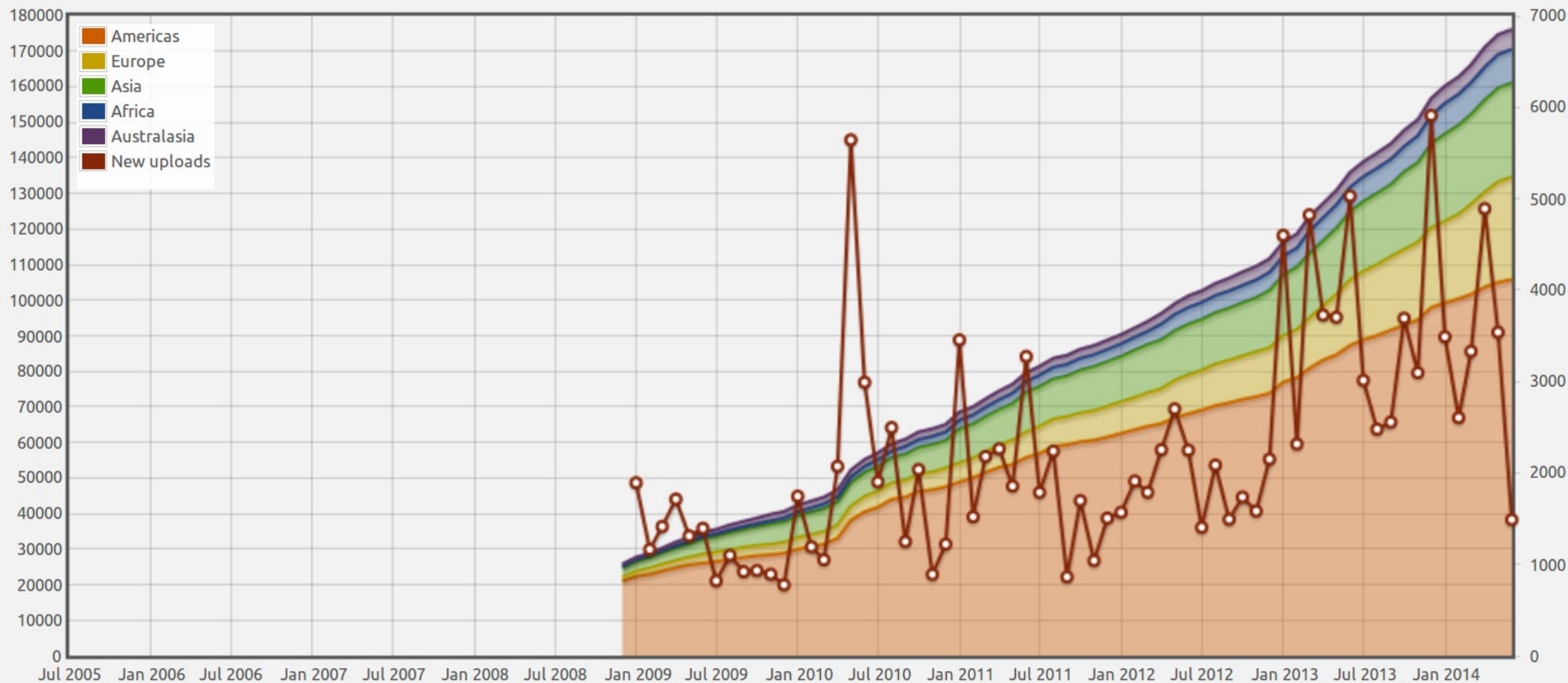


Xeno Canto corpus

Large scale Bird database

load

Go



Welcome to Bird50 : 50 species extracted from thousands
a subset of the Life Clef 2014, 2015, 2016 challenge

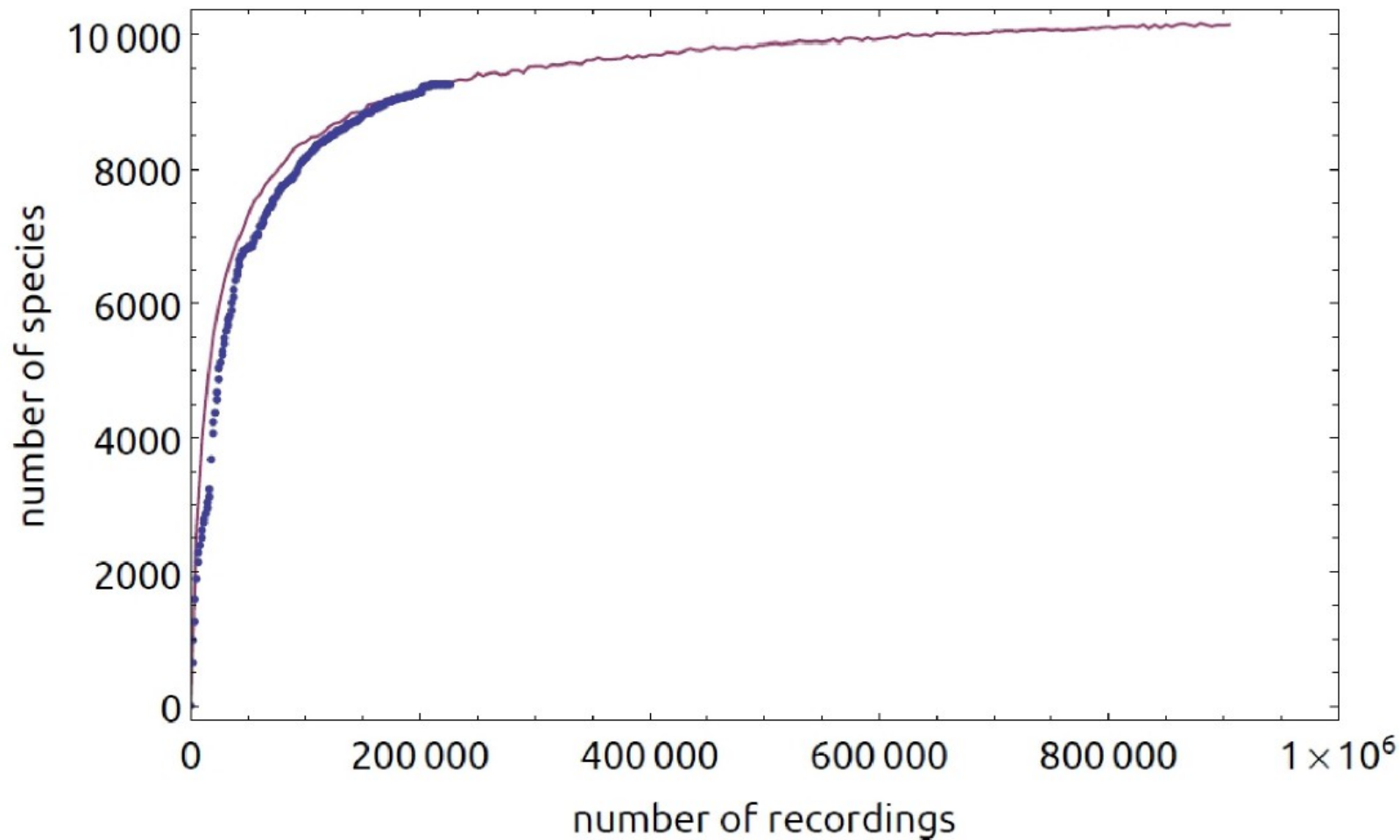
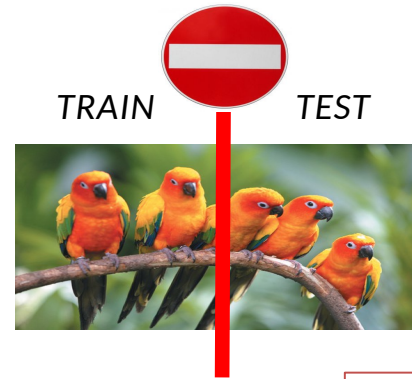


Fig. 8. Extrapolation of species accumulation curve assuming that as yet unrepresented species have a probability of being uploaded that is $1/3$ that of a species represented with 1 recording. It is likely the collection needs to multiply in size before completion at the species level is reached.

[From Xeno Canto Clef 2015 paper]

BirdCLEF: Task

**Event-based split (2/3 vs. 1/3)
Foreground & background species**

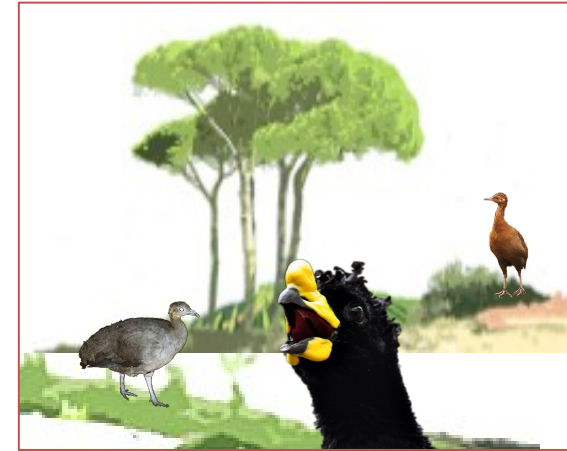


*same day,
same time,
same place,
same device,
same species*

Numerous audio records are associated with some background species in the metadata

But most of them focus on one single bird recorded with a directional microphone

 **Training Set**



Bg sp. 1

Bg sp. 2

Foreground sp.

Test recording 2

Test recording 238

Test recording 490

2; xirfzd;1;0.23009059
2; onzvd;2;0.20605771
 2; ziecey;3;0.19009507
 2; asvgnp;4;0.15429588
 2; Jathqge;5;0.10588206
2; gmeajz;6;0.077025205
2; zxtbno;7;0.036553476
 238; dkecip;1;1.0
 490; asvgnp;1;0.47851267
 490; athqge;2;0.24759477
 490; onzvd;3;0.158397
 490; zxtbno;4;0.06804684
 490; xirfzd;5;0.047448702

**Metric = Mean
Average Precision**

A difficult challenge :

From the 137 registred teams, only 6 submitted a run ...

China Ac. Science Beijin

Golem Mexico

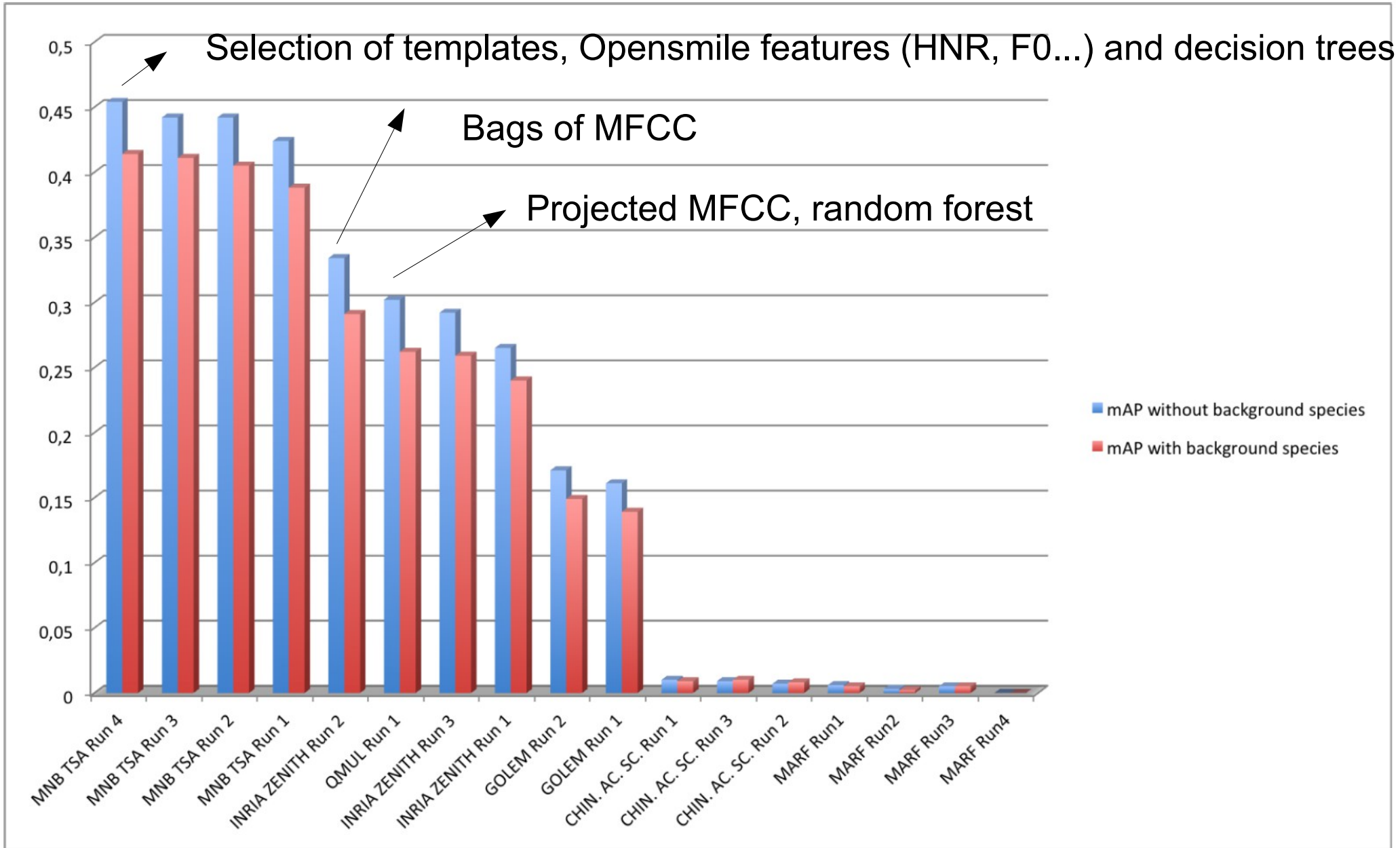
INRIA Zenith

Queen MU London

Berlin Museum Natural Science

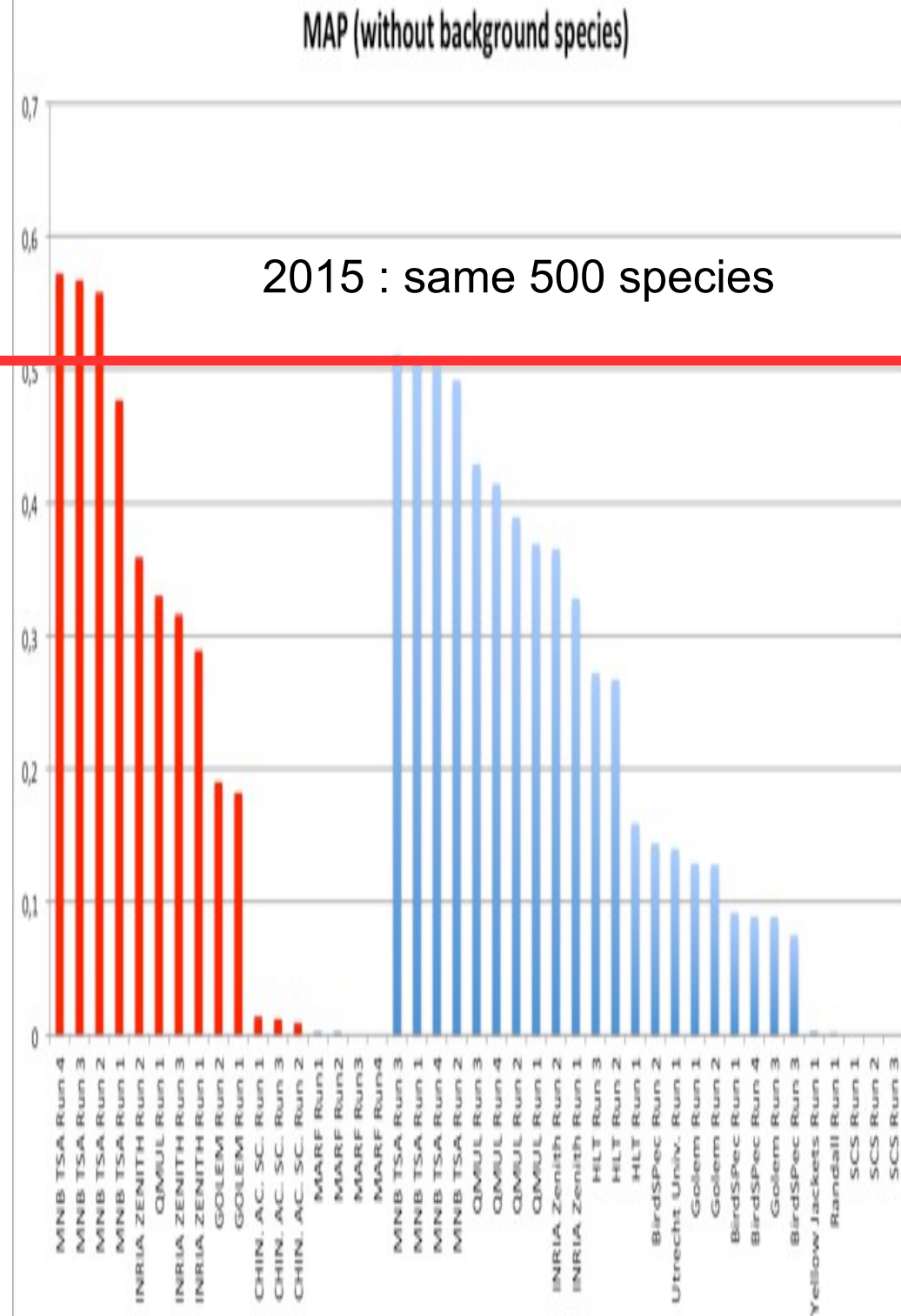
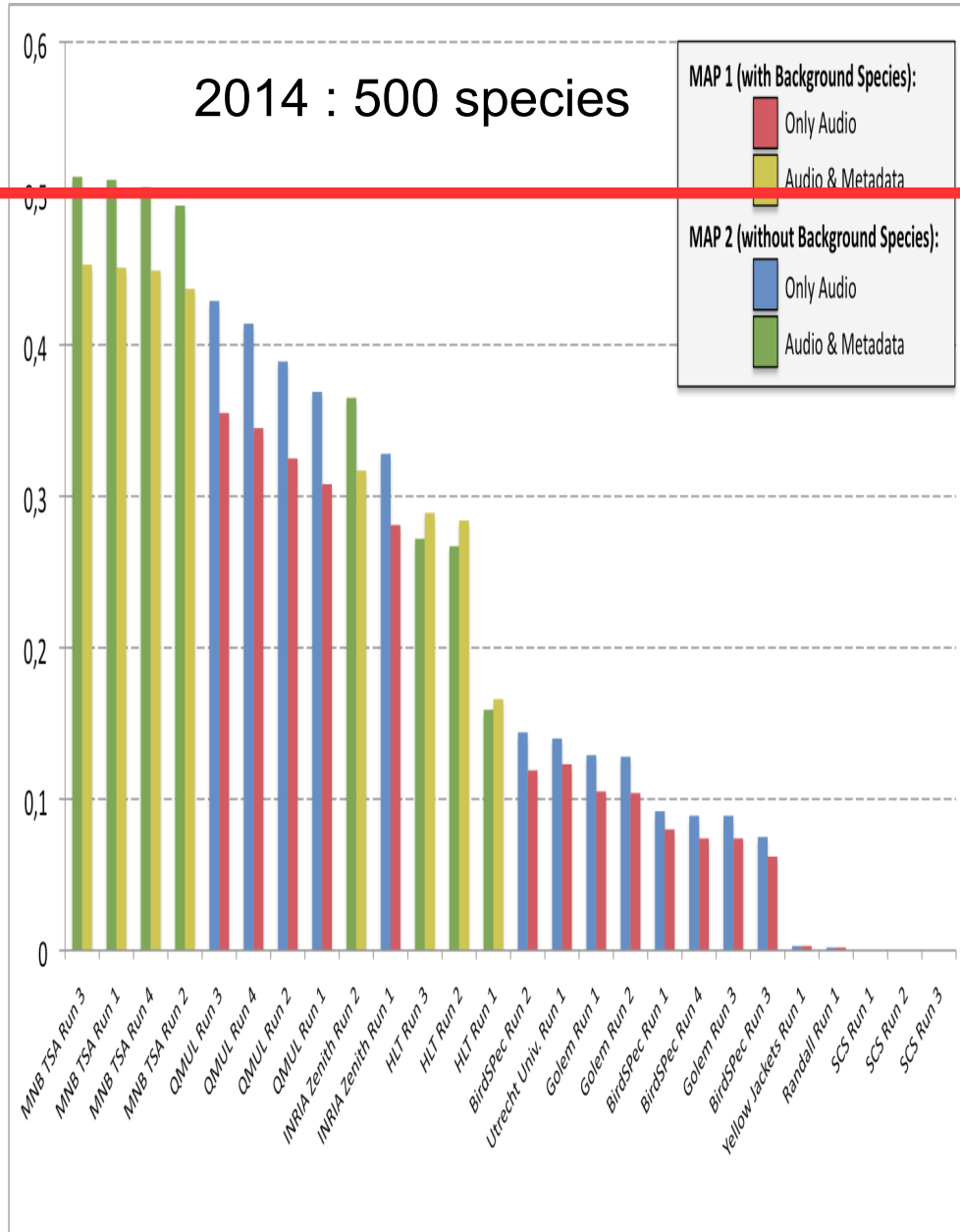
MARF Canada

MAP results : only foreground species (blue) and background species (red)



500 species MAP : 2014 vs 2015

Demonstrate small improvement ...



Corpus online =

- http://sis.univ-tln.fr/~glotin/SABIOD_LIFECLEF/BIRD50/

AmazonBird50_training_input.tar.gz 2,4 Go (924 files)

AmazonBird50_testing_input.tar.gz 1 Go (375 files)

AmazonBird50_training_output.csv with 1 label per file [1,50]

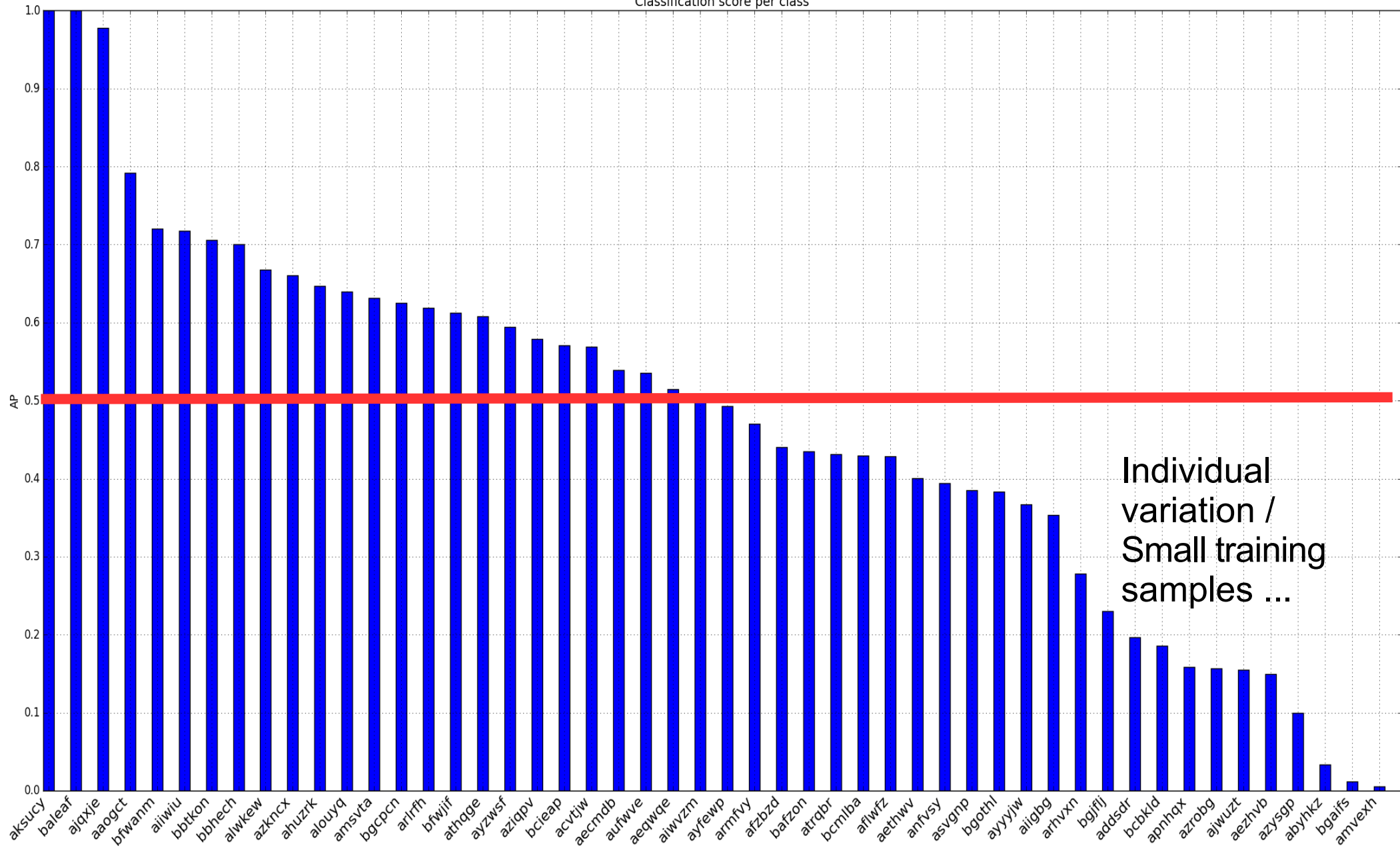
AmazonBird50_testing_output.csv with 1 label per file [1,50]

- Metrics = Precision = Average Precision, MAP over the 50 species

Variation of the Average Precision per Species

Automatic classification on 50 species of SABIOD bird50 (DYNI LSIS test)

Classification score per class



[Home](#)

[ICML 2014 Workshop](#)

[NIPS 2013 Workshop](#)

[ICML 2013 Workshop](#)

[Data Samples](#)

[Online Cetacean Tracking](#)

[Team](#)

[Media](#)

[Contact](#)

Bioacoustic signaling is a primary mode of communication and exploration for most of the animals. It enables quick load and transfer of information without any visible contact with the target, tackling the reduced visibility of deep forest (insect, frogs, birds, mammals...), cave or night activities (insects, bats), and/or the long distances like in ocean (krill, fishes, whales...). Bioacoustics is also one of the factors in optimizing natural selection, playing a significant role in signalling resource qualities to potential partners. The SABIOD project aims to detect, cluster, classify and index bioacoustic big data in various ecosystems, at different space and time scales, in order to reveal informations on the complex sensori-motor loop, and on the health of an ecosystem, yielding to new biodiversity insights.

NEWS:

- [ICML 2014 workshop opens : We are pleased to inform you that paper submission is now open for our ICML 2014 workshop uLearnBio: "Unsupervised Learning from Bioacoustic Big Data" joint to ICML, 21-26 june 2014 in Beijing, China.](#)

- [LifeCLEF 2014 Bird Identification Task - Training Data Release](#) We invite you to participate in this new Bird Identification challenge, based on different types of audio records over 500 species from South America centered on Brazil (14027 audio records, 9688 train+4339 test). This challenge is the biggest ever organized, after the pionner ICML 2013 and NIPS 2013 bird challenges that dealt 'only' with less than one hundred of species but interested many teams. [Listen here to a sample of the LifeClef challenge.](#)

- [Book of the int. workshop Neural Information Scaled to Bioacoustics NIPS4B, from neurons to big data - Nevada](#)

- [Book of the 1st international Mastodons Big Data workshop for bioacoustics ICML 12 - Atlanta](#)

Neural Information Processing Scaled for Bioacoustics -from Neurons to Big Data-

Proceedings of NIPS4B, international workshop joint to NIPS, USA, 2013

Glotin H., LeCun Y., Artières T., Mallat S., Tchermichovski O., Halkias X.
Toulon, New-York, Paris

