

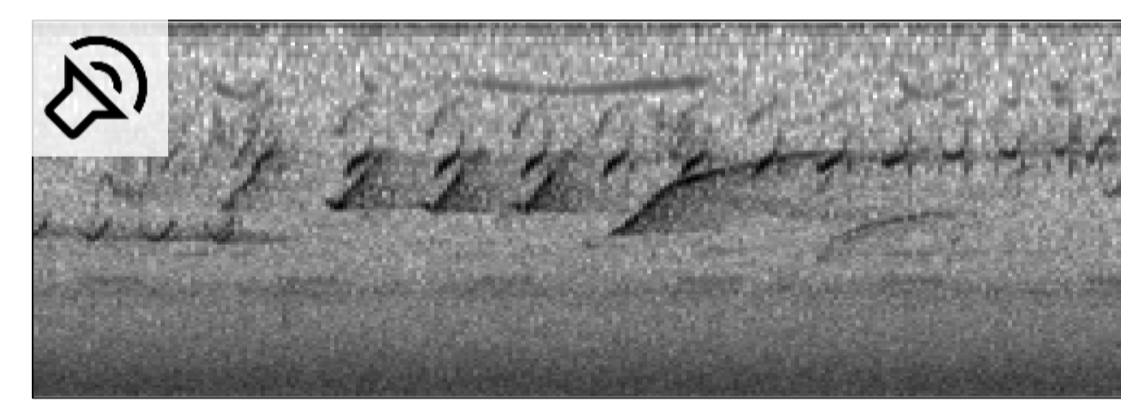
MetaPerch: Learning from Metadata for Bioacoustics Foundation Models

Mustafa Chasmai[■], Vincent Dumoulin, Jenny Hamer[◆]
[■]UMass Amherst; work done as a student researcher at Google DeepMind
[◆]For correspondence: hamer@google.com

If you can only spare a glance 👁️

- **Bioacoustics** is the study of nature through sounds produced by living things. AI and ML have a massive impact potential there.
- Foundation model training in this field relies on **large scale citizen science platforms**. The rich **metadata** these platforms contain is often ignored for training.
- We present a **thorough, systematic exploration** of the use of metadata—such as location and time—as auxiliary supervision signals.
- A valuable byproduct of this exploration is **MetaPerch**, a new bioacoustic foundation model that uses metadata as auxiliary targets.

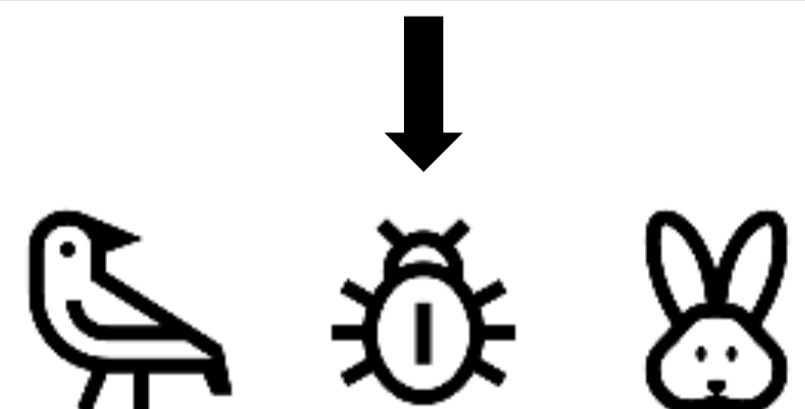
Introduction



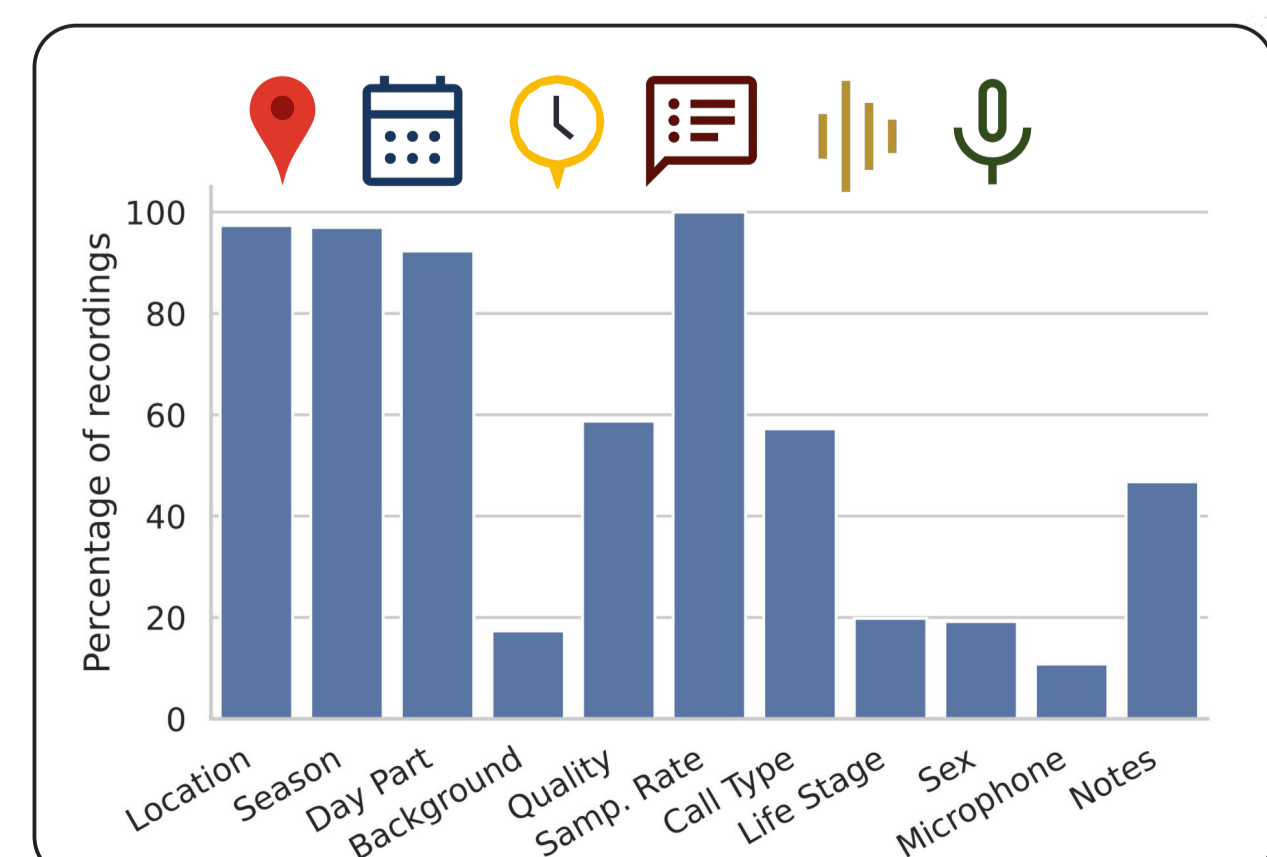
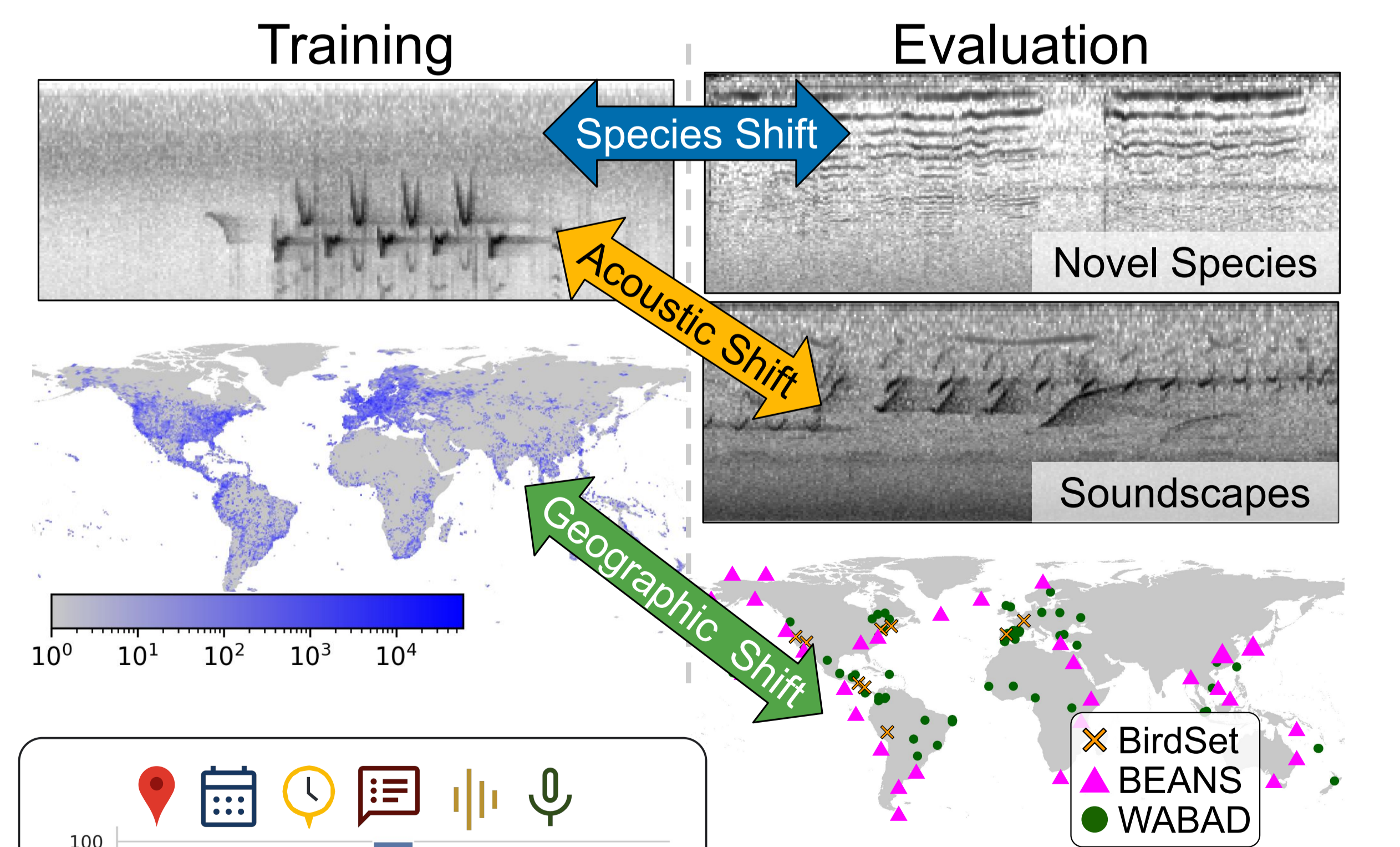
Task: Identify all species vocalizing in a given audio recording

Applications: Biodiversity monitoring & conservation

Challenges: Fine-grained categories, data scarcity



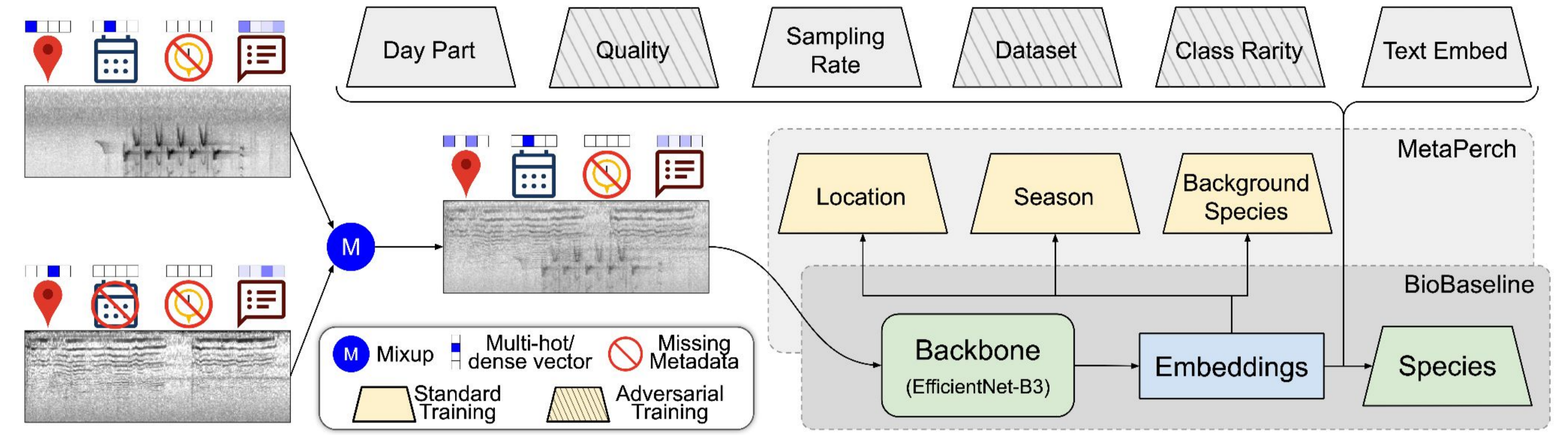
Motivation



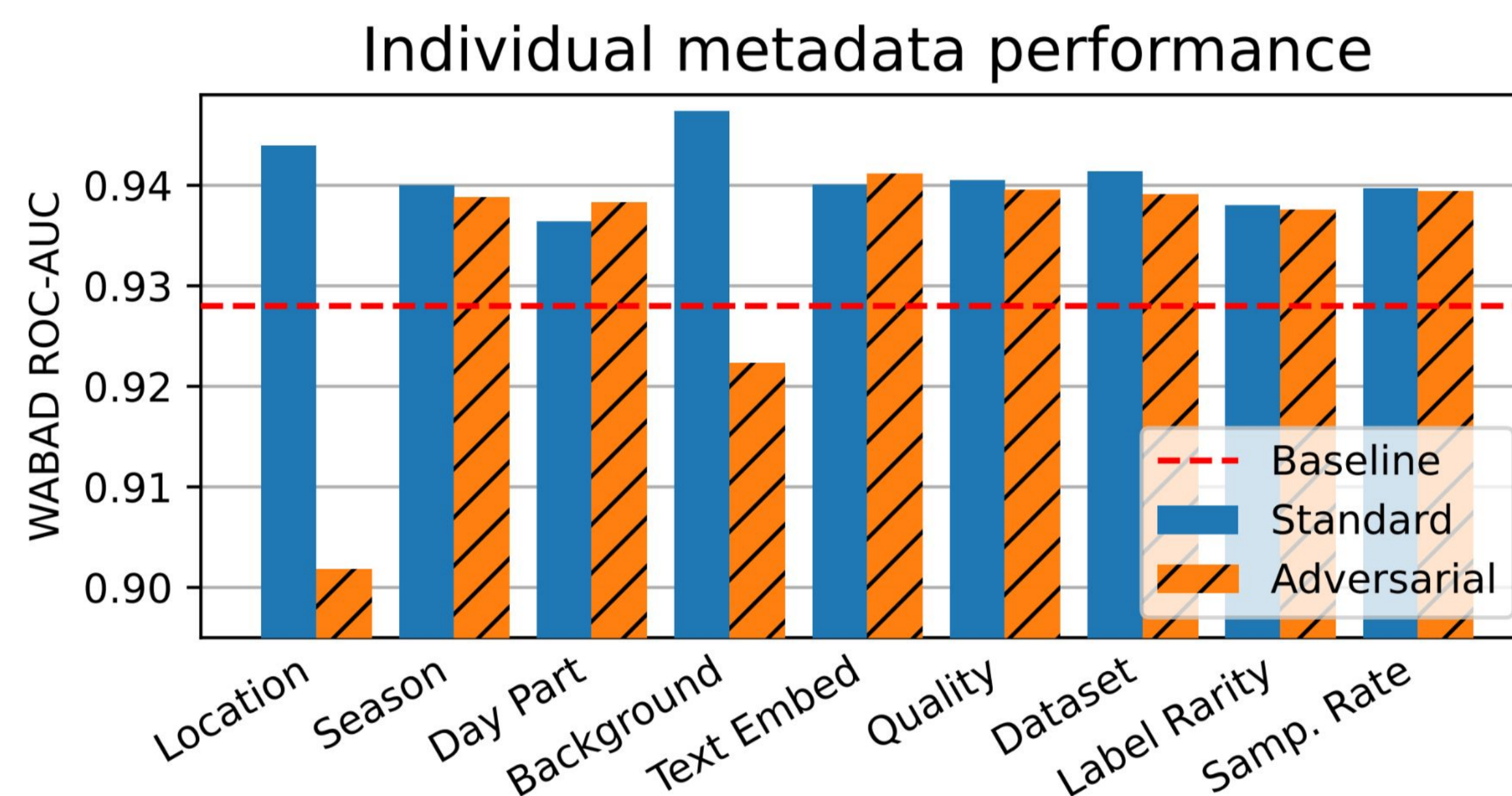
- Metadata**
- Available only at training time
 - Correlated with species, but not expensive to annotate
 - Can help learn embeddings that generalize better to domain shifts

MetaPerch

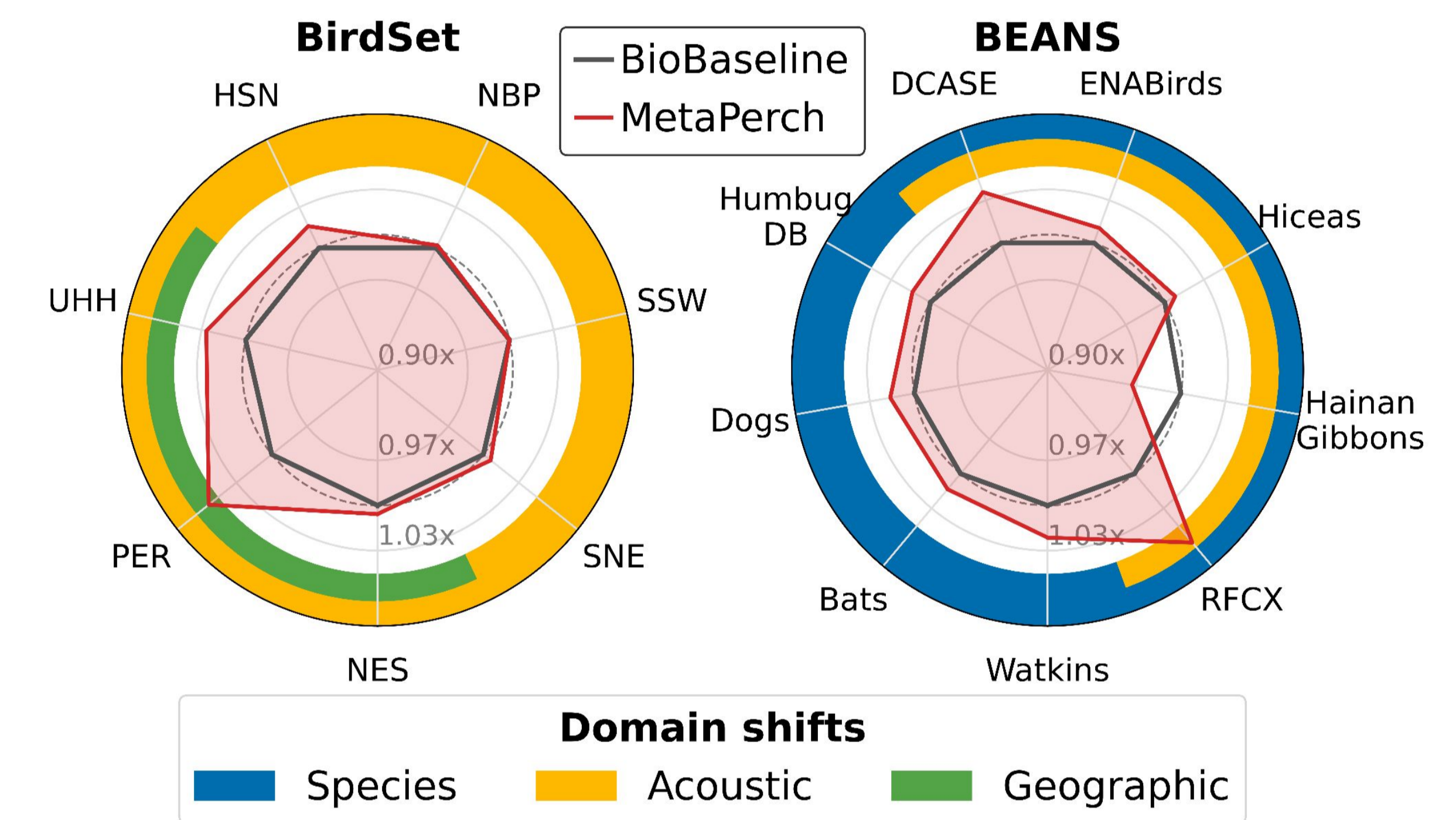
- Auxiliary training tasks: predict each metadata solely from audio
- Each metadata can be used for standard or adversarial training (for robust and spurious metadata)
- No use of metadata at deployment



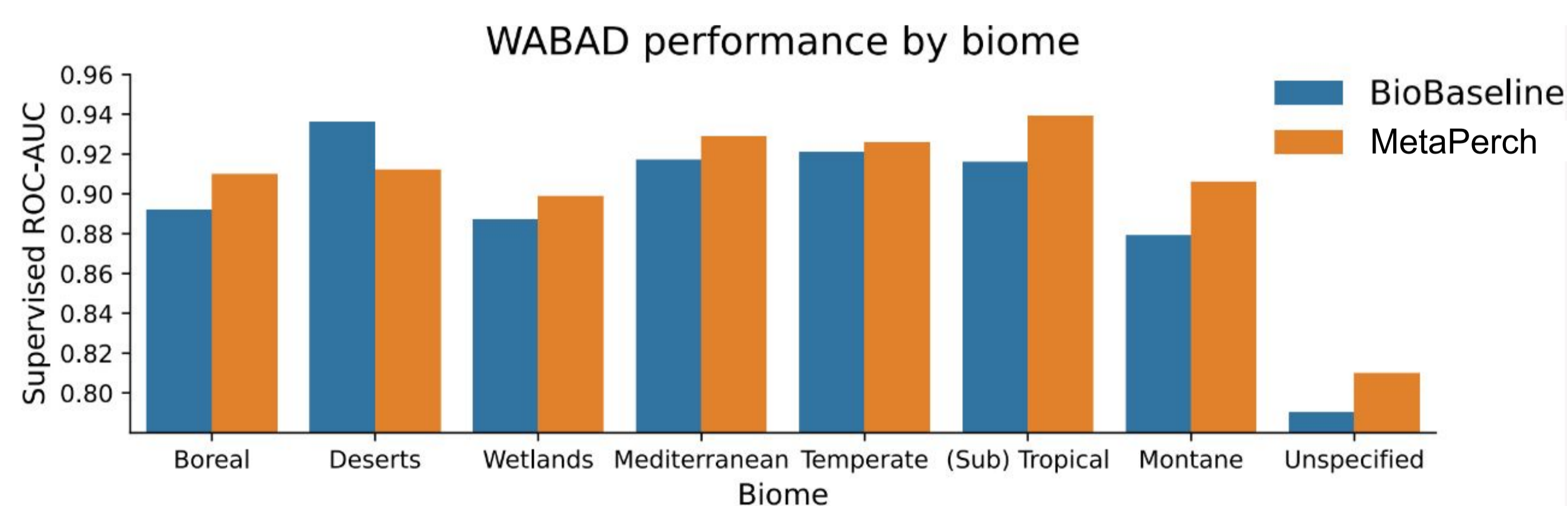
Performance on bioacoustic benchmarks



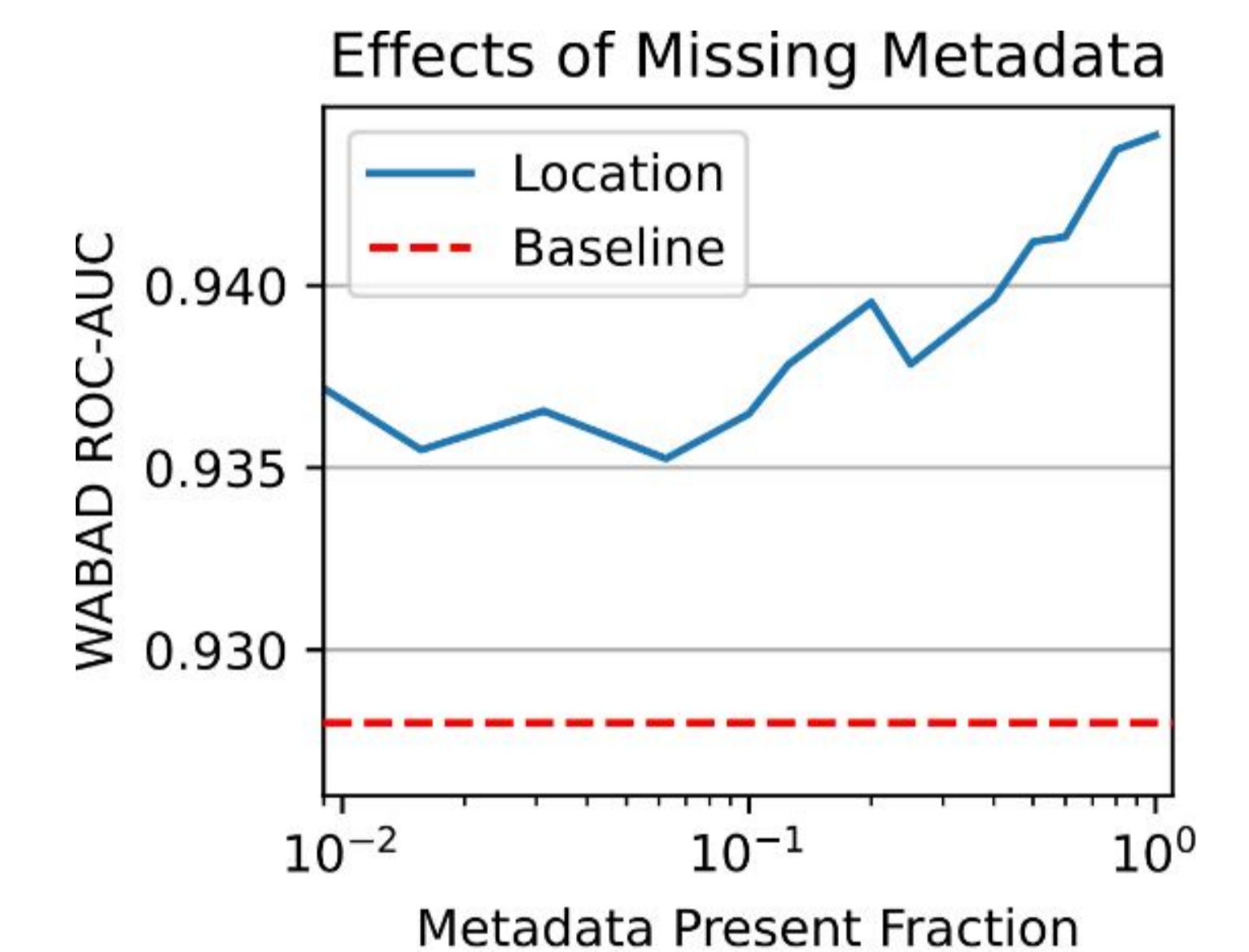
Performance of MetaPerch when trained using individual metadata, using standard and adversarial loss formation



Performance on datasets grouped by respective domain shifts



Performance by biome on WABAD, a global bird sounds dataset



MetaPerch performance in the case of missing metadata

Empirical findings 🔍

- BirdSet**
- Metadata **improves transfer to soundscapes** for known species.
 - Metadata **aids performance in underrepresented regions**.
- BEANS**
- Metadata **helps transfer to novel species** for focal recordings.
 - Metadata **helps transfer to soundscapes** for novel species.
- WABAD**
- Metadata **helps in globally-distributed tasks**.
 - Metadata **benefits vary across biomes**.

Discover more ✨

Read the paper!



Use the model!

