Explainable Deep Learning and Approximate Bayesian Computation for population genomics inference

Brazier, Thomas[1] (contact: thomas.brazier@univ-rennes.fr)

1 University of Rennes, CNRS, ECOBIO (Ecosystems, Biodiversity, Evolution) - UMR 6553, Rennes, France

Approximate Bayesian Computation (ABC) is a popular likelihood-free inference method for parameter inference in biology, especially in population genetics.

Limitations
• High-dimensional sets of summary statistics (curse of dimensionality)
• Evaluating model complexity (model misspecification)
• Model validation & interpretability
• Explainability

How to address the emergence of large genomic datasets and increase in model complexity (*msprime, SLiM*)?
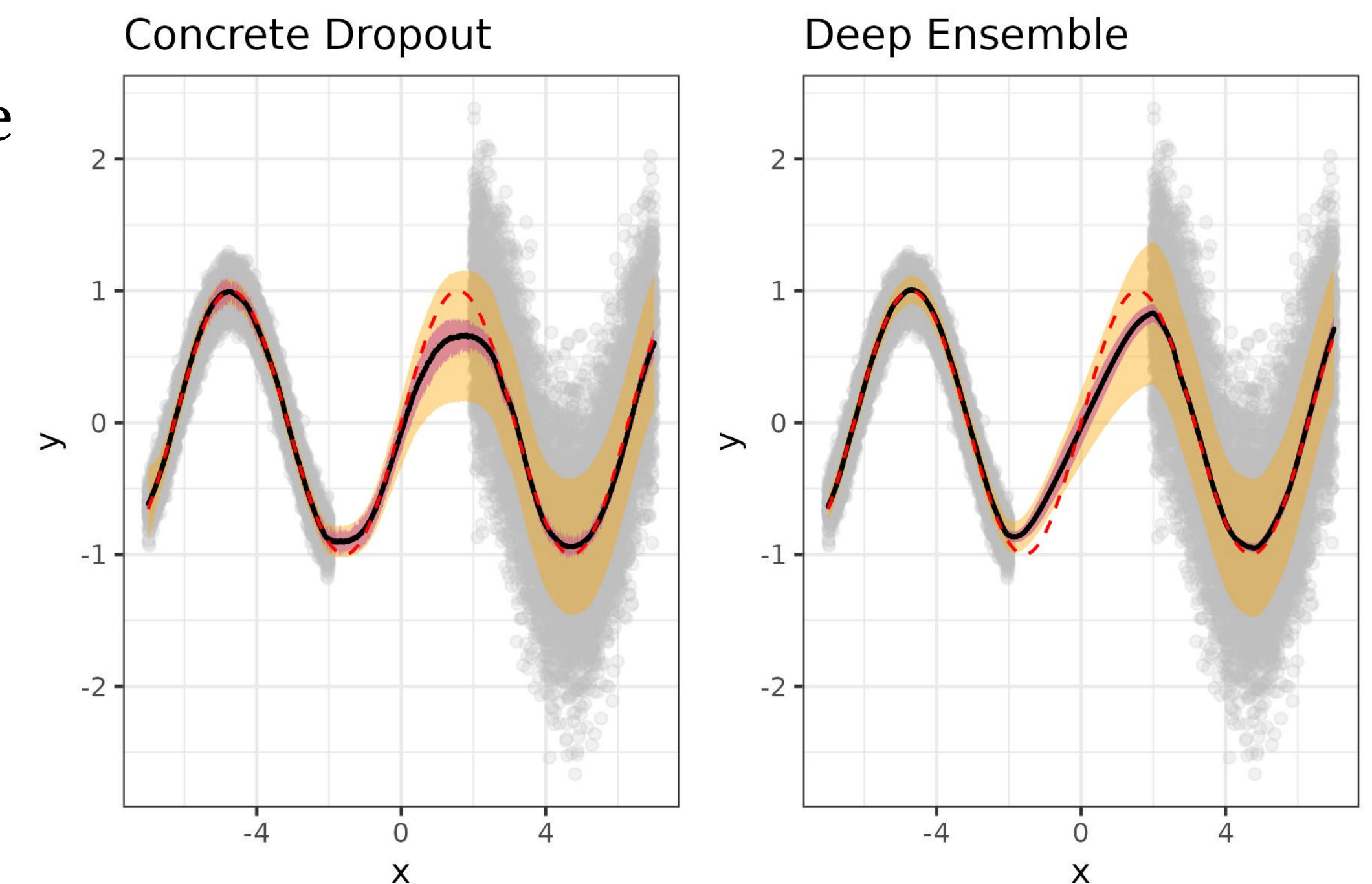

Fig. 1. Aleatoric (orange) and epistemic (purple) uncertainty for model validation and prediction confidence, for Concrete Dropout (Gal et al. 2017) and Deep Ensemble (Lakshminarayanan et al. 2017). Training data in light grey. Expected in dashed red.

ABCNeuralNet. A R package for Bayesian Deep Learning

4 Neural Net methods
• Concrete Dropout
  (Gal et al. 2017)
• Deep Ensemble
  (Lakshminarayanan et al. 2017)
• Tabnet-ABC (*NEW*)
  (Arik et al. 2021)
  (Åkesson et al. 2021)
• Monte Carlo Dropout
  (Gal 2016)

(1) **Bayesian Neural Networks** implemented in R ***torch***.

(2) Fit the model (simulations and summary statistics) with *luz*.

(3) Model diagnostic & tuning (performance) and evaluate model misspecification with test data and **aleatoric + epistemic uncertainty.**

(4) Calibrated Credible Interval with **Conformal Prediction** (Baragatti et al. 2024).

(5) Explain predictions with 10 methods of **Feature importance** (*innsight* R package).

A user-friendly package, modern stack and end-to-end workflow (diagnostic plots and methods).

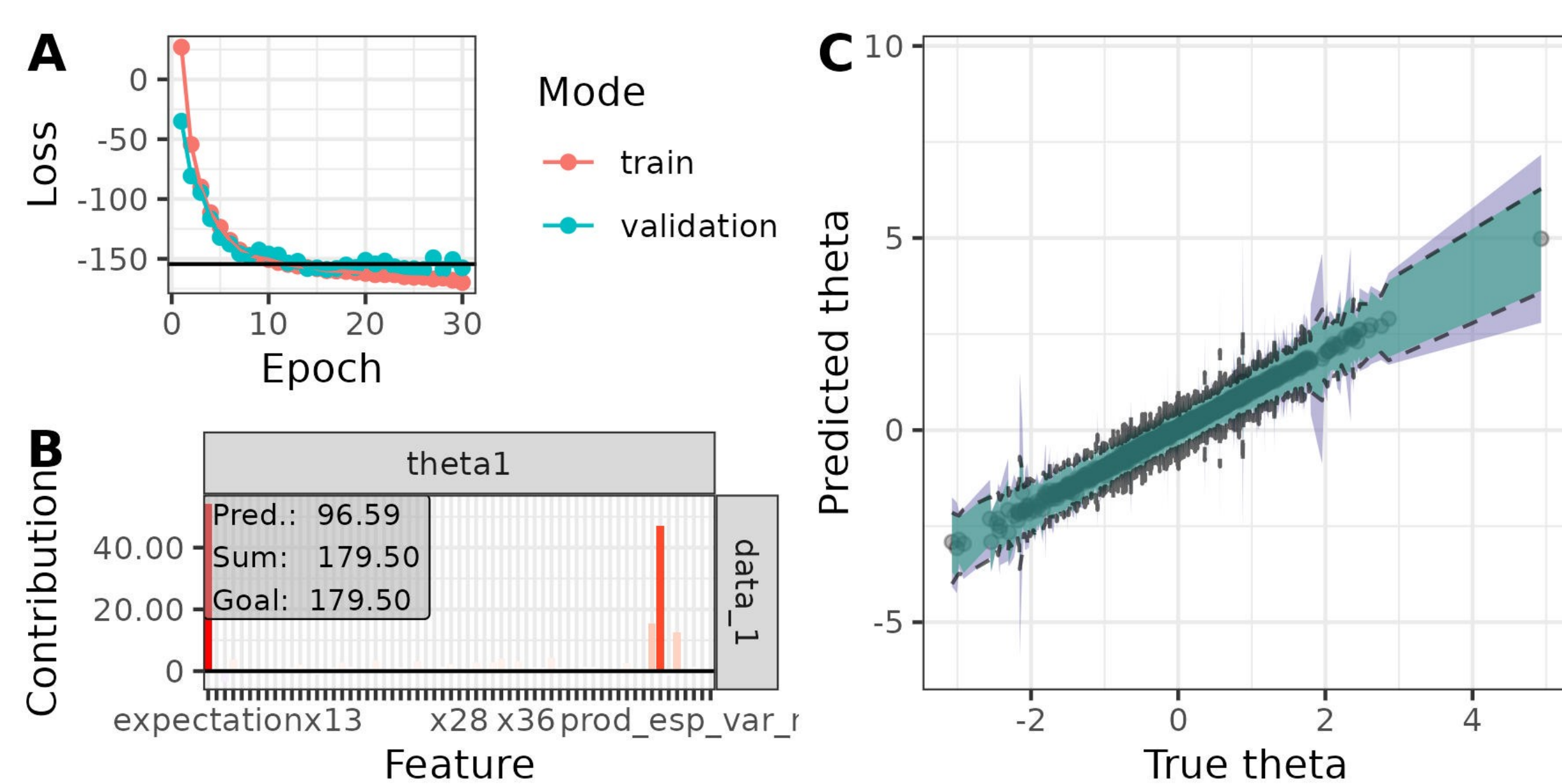## Workflow & performance on a toy model

Fig. 2. Workflow of fitting, checking and explaining the model with ABCNeuralNet with Concrete Dropout on 100,000 training samples. (A) Training curve. (B) Feature Importance with 'DeepLift'. 50 of 61 variables are random noise. (C) Predictions as a function of true simulated value (1,000 test samples) with Conformal Prediction Credible Intervals expected (true value, dashed line) and estimated from epistemic (purple) and overall uncertainty (green).
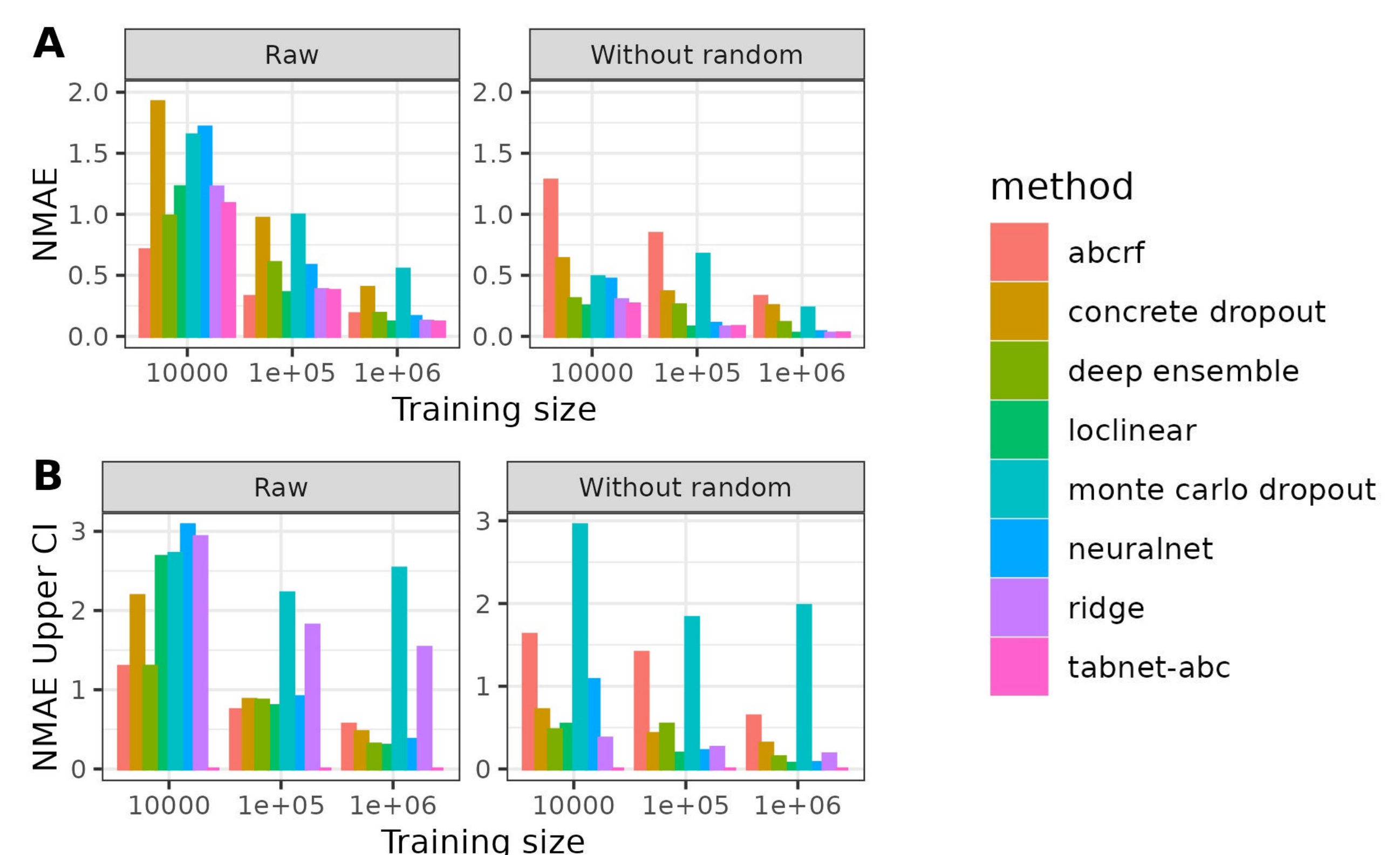

Fig. 3. Performance of ABCNeuralNet compared to other ABC reference methods. Raw dataset with 50 random noise variables (left) and dataset without noise variables (right). (A) NMAE between predictions and true expected value. (B) NMAE between the 95 % Credible Interval estimated and true C.I.

... and on a realistic human population genomic dataset?

Fit/training   Evaluate (1000 test samples)   Explain (deepSHAP)   Predictions with C.I. (observed)
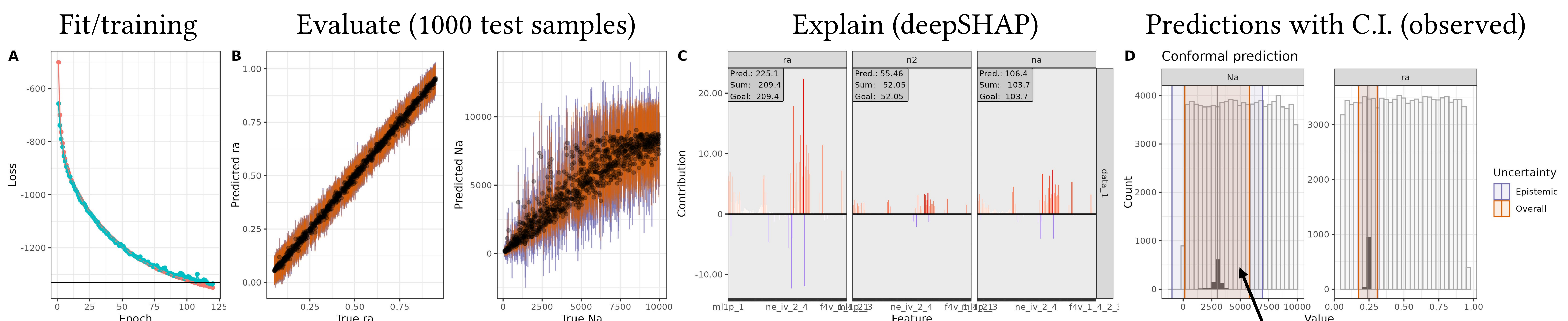
Neural net performance, metrics

When the model don't know