Bayesian Methods for Enhanced Greenhouse Gas Emissions Inventories

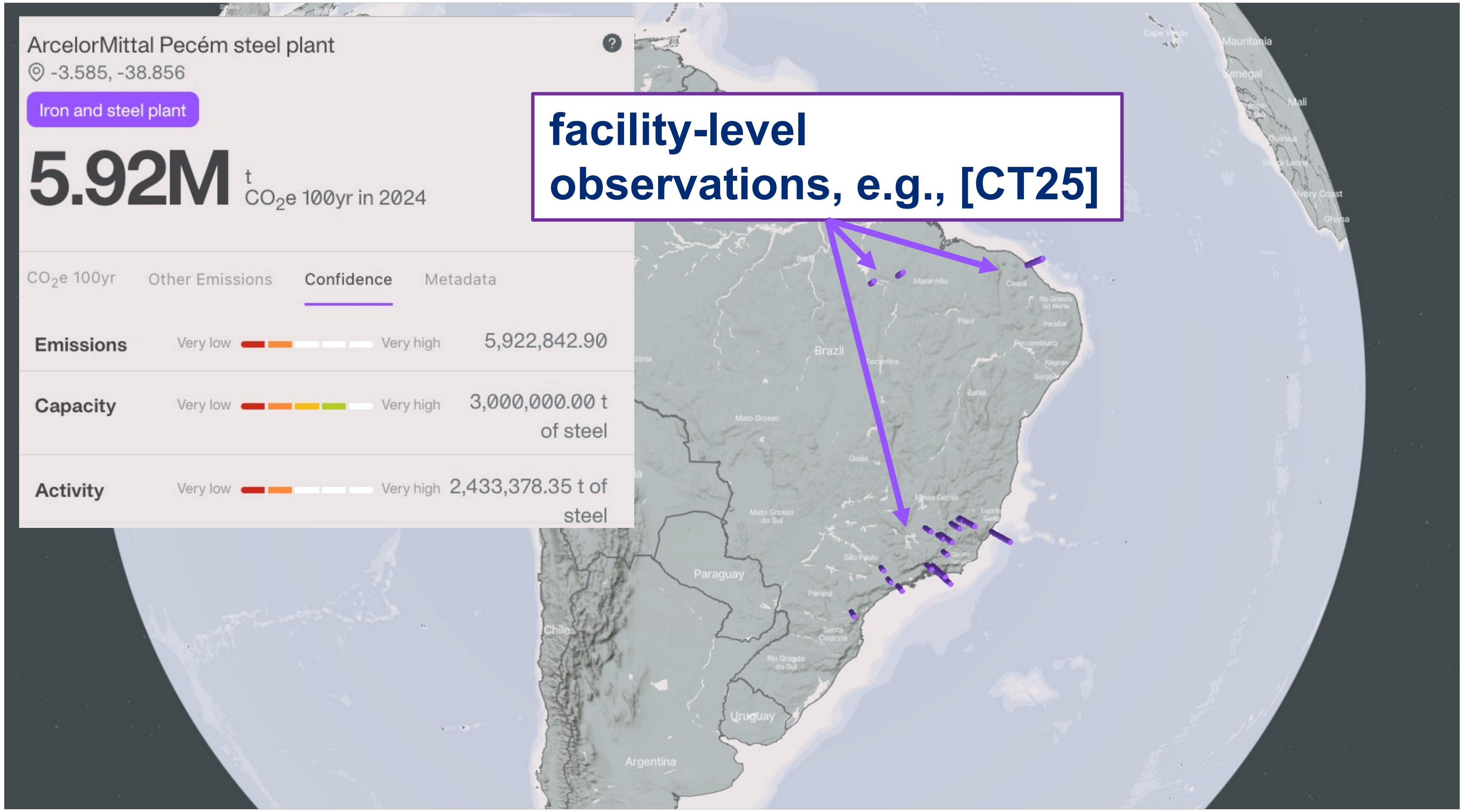
Tackling Climate Change with Machine Learning workshop at NeurIPS 2025

Michael Pekala[1,3], Zoheyr Doctor[2,3], Elizabeth Reilly[1,3], Gavin McCormick[2,3], Gary Collins[1,3], Daniel Moore[2,3], Amy Piscopo[2,3], Krsna Raniga[2,3], Michael Robinette[1,3], Ting So[2,3]

[1] Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland, USA,
[2] WattTime, Oakland, USA,
[3] Climate TRACE

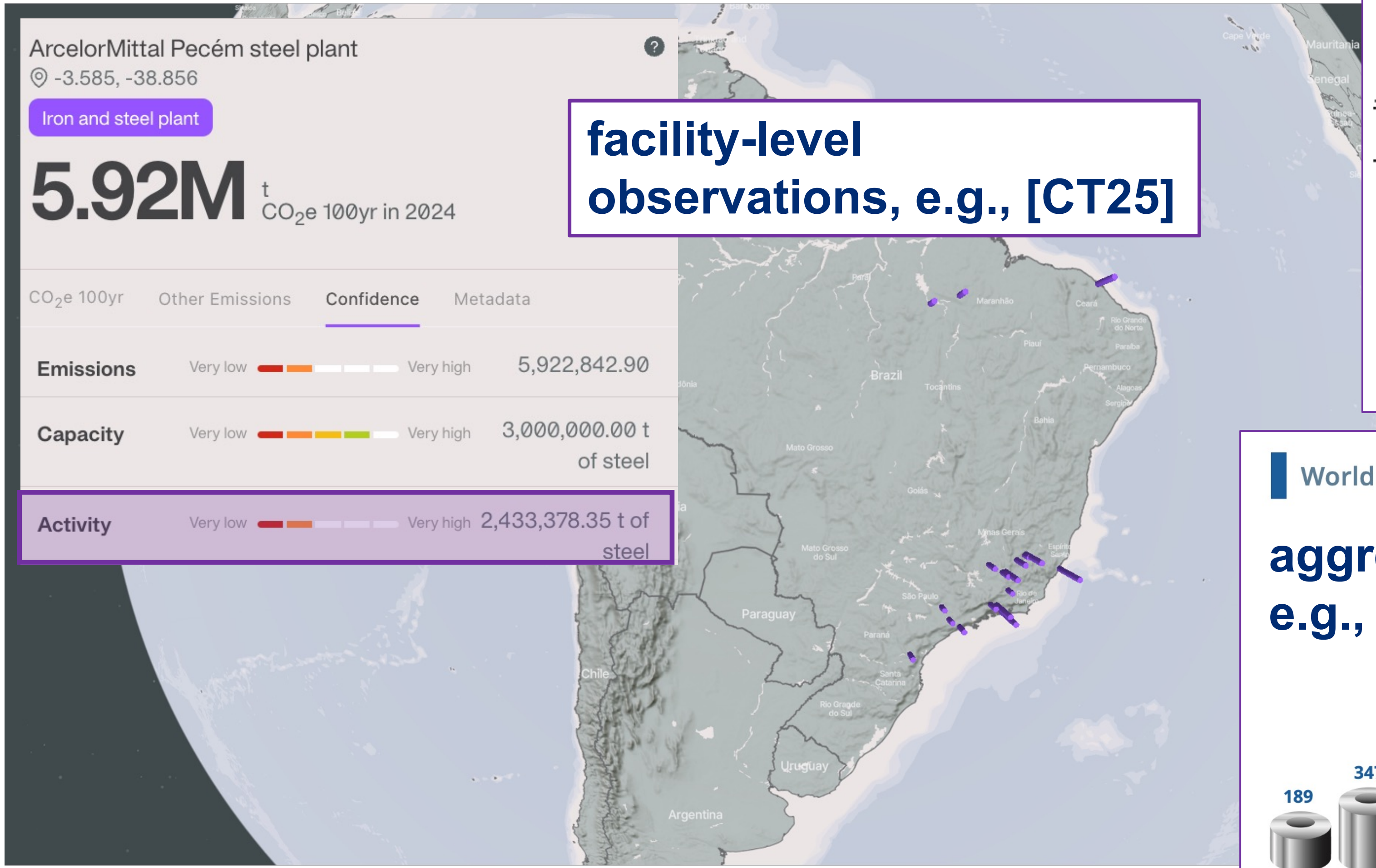
Contact: michael.pekala@jhuapl.edu, elizabeth.reilly@jhuapl.edu, or zoheyr.doctor@watttime.org

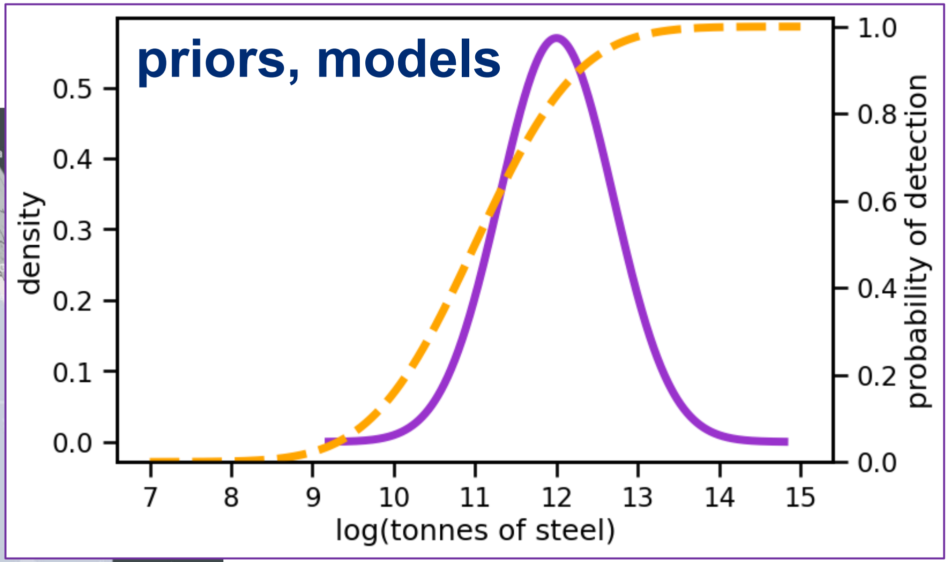
What emissions properties can we estimate for entities we don’t directly observe?

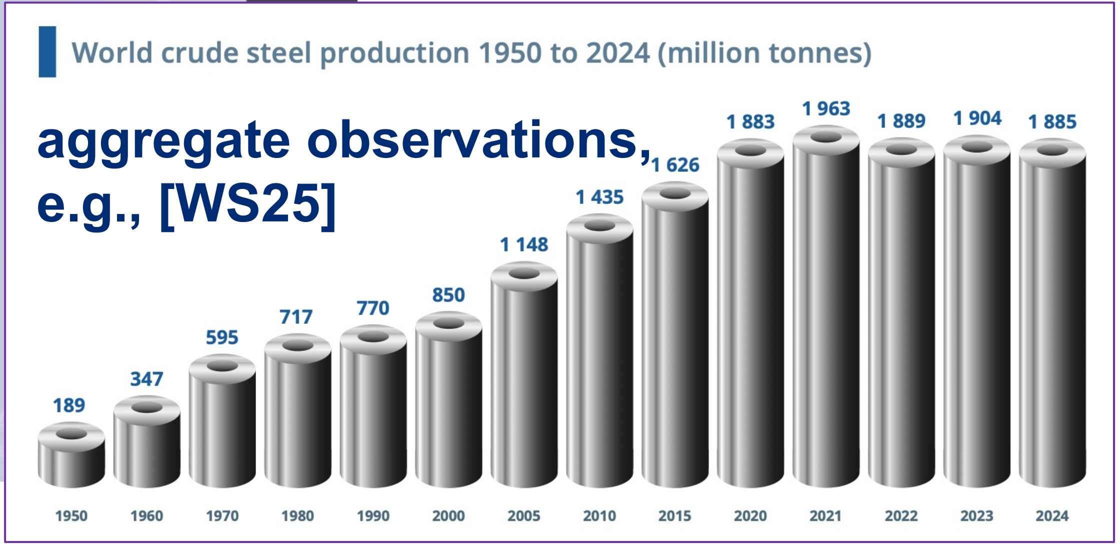


[CT25] Climate TRACE website https://www.climatetrace.org

What emissions properties can we estimate for entities we don't directly observe?







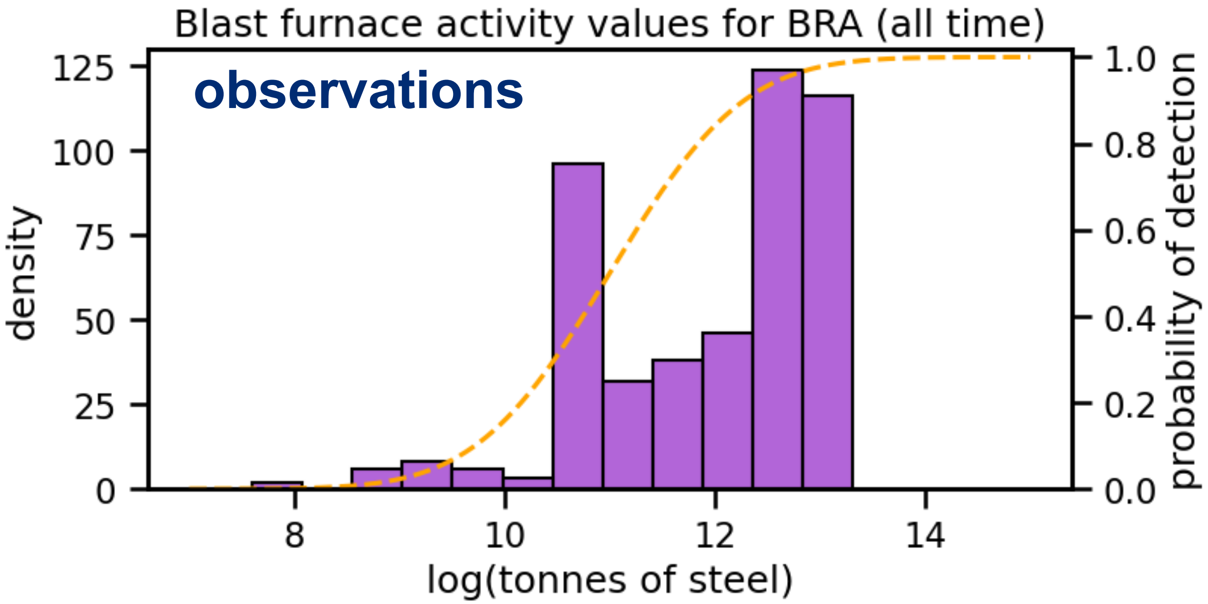
[CT25] Climate TRACE website https://www.climatetrace.org
[WS25] World Steel Association. World Steel in Figures, 2025. https://worldsteel.org/wp-content/uploads/WSIF-2025-infographic.pdf

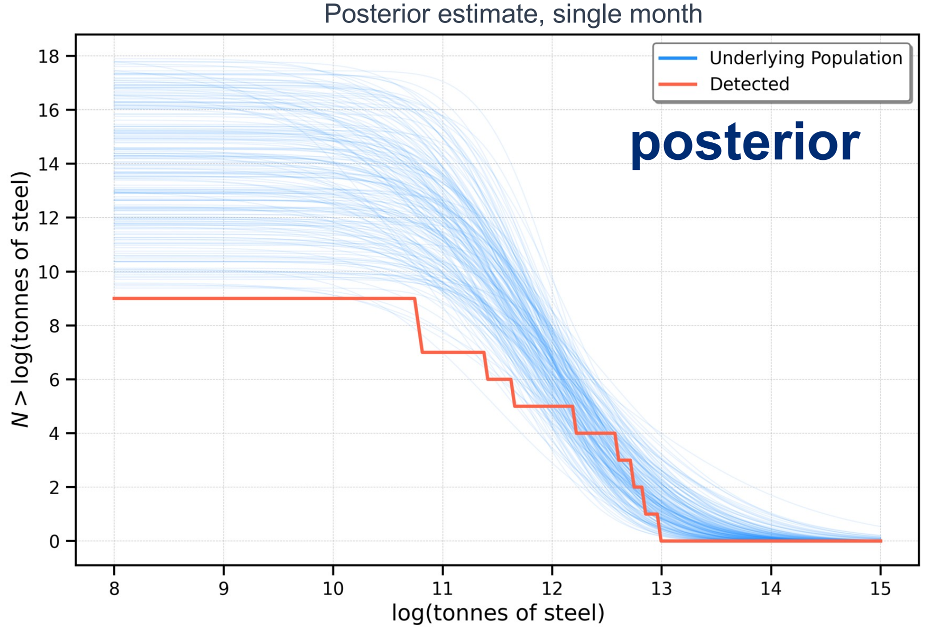
Parameter estimation amid selection bias

Following the hierarchical Bayesian approach of [Man19]:

$$\boxed{p(\lambda, N | \{d_i\})} = \pi(\lambda)\pi(N) \prod_{i=1}^{N_{\mathrm{obs}}} \frac{\int p(d_i|\theta) p_{\mathrm{pop}}(\theta|\lambda) d\theta}{\int p_{\mathrm{det}}(\theta) p_{\mathrm{pop}}(\theta|\lambda) d\theta} \cdot e^{-N_{\mathrm{det}}} \left(N_{\mathrm{det}}\right)^{N_{\mathrm{obs}}}$$

$$d_{\mathrm{net}} \sim \mathcal{N}(N \cdot \mathbb{E}[p_{\mathrm{pop}}(\theta|\lambda)], \sigma_{\mathrm{net}})$$





[Man19] Mandel, Ilya, Will M. Farr, and Jonathan R. Gair. "Extracting distribution parameters from multiple uncertain observations with selection biases." MNRAS 486.1 (2019): 1086-1093.
[Pha19] Phan, Du, Neeraj Pradhan, and Martin Jankowiak. "Composable effects for flexible and accelerated probabilistic programming in NumPyro." arXiv preprint arXiv:1912.11554(2019).

Conclusions & next steps

- Early indications suggest the model does a good job capturing parameters of interest in synthetic data
- Furthermore, the distributional assessments provide insight into how active these facilities might be
- Next steps include:
  - explore other distributions for modeling facility populations (e.g., power law, uniform); run sensitivity studies, cross-validation, and model selection
  - temporal extensions (move beyond just a single month)
  - more complete study of detection models and noise priors
- Please see our CCAI 2025 poster for more details

Contact: michael.pekala@jhuapl.edu, elizabeth.reilly@jhuapl.edu, or zoheyr.doctor@watttime.org

The authors would like to thank the Climate TRACE coalition for its organizational support, and Climate TRACE's funders for their financial support. For a full list of funders, please visit the Climate TRACE website: https://climatetrace.org/team

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY