Machine Learning Models with Uncertainty Quantification for Precipitation Type Prediction

Dhamma Kimpara*, Belen Saavedra
Charlie Becker, David John Gagne, Gabrielle Gantos, John Schreck

*NCAR **M**achine **I**ntegration and **L**earning of **E**arth **S**ystems*
**University of Colorado Boulder*

NCAR

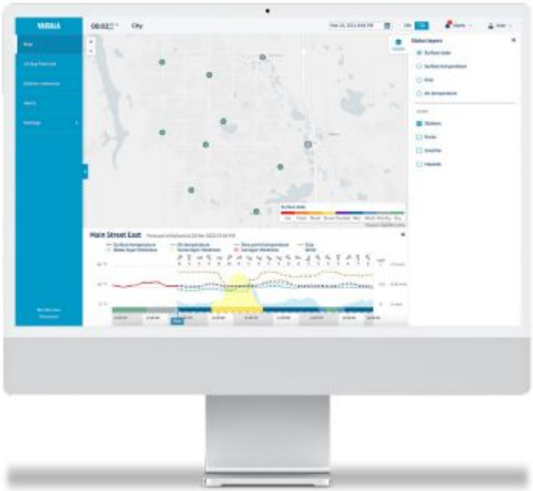
August 2, 2023

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Precipitation type greatly affects impact of winter storms

Vaisala Wx Horizon Pro

Bring actionable insights and predictions to your winter road maintenance

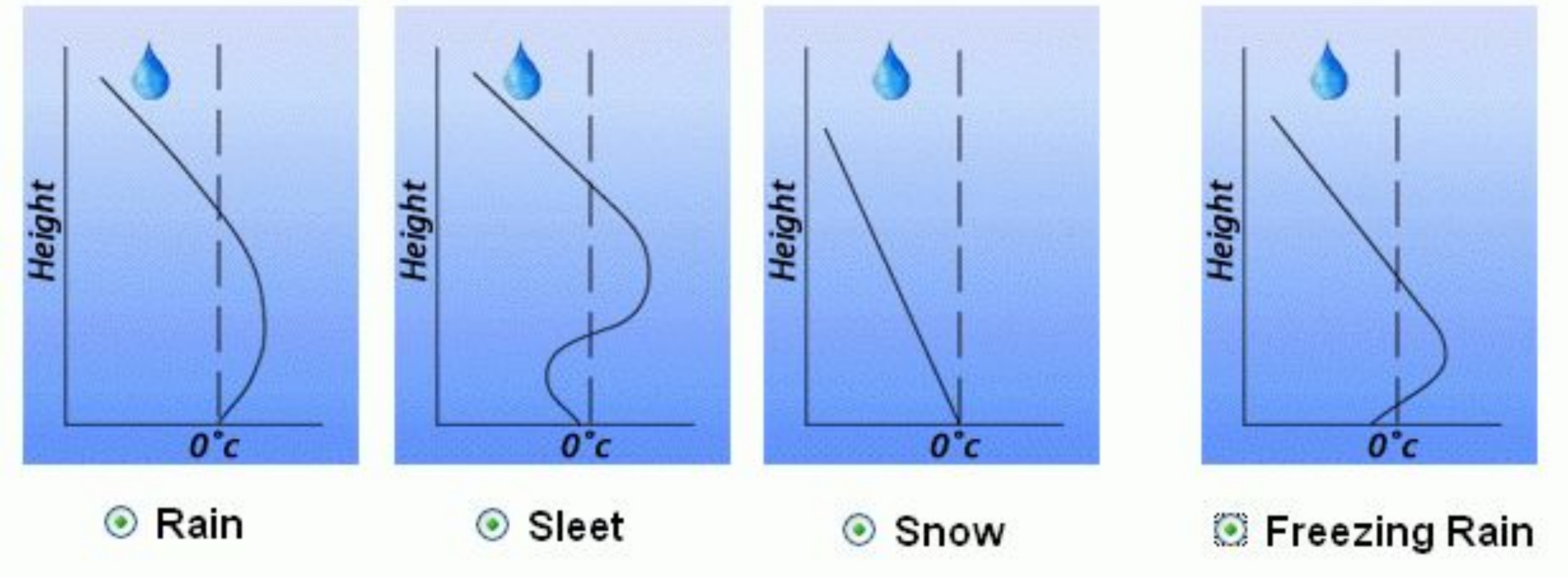


What did I do?

- Evaluated precipitation type models
  - map-reduce computations on large Xarrays (i.e. banging my head against the cluster)
- Extended Machine Learning Methods for Uncertainty Quantification
- Wrote a small utility for submitting PBS jobs in python (https://github.com/dkimpara/pbs_utils)

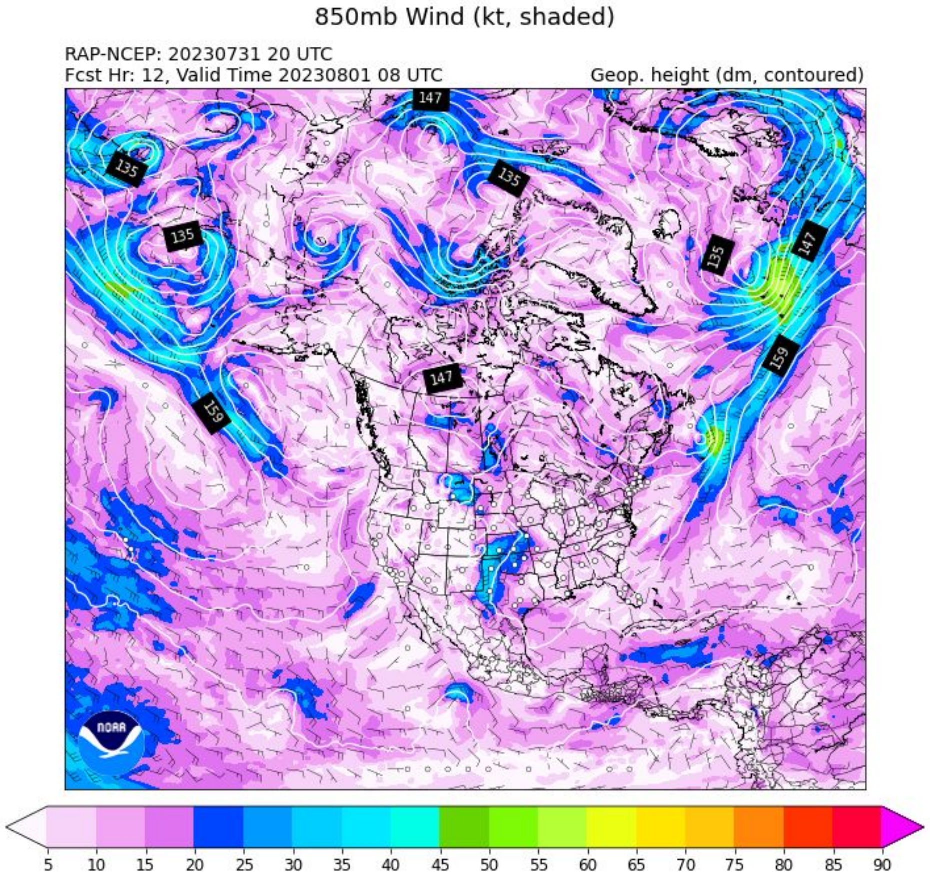
How do we predict precipitation type?

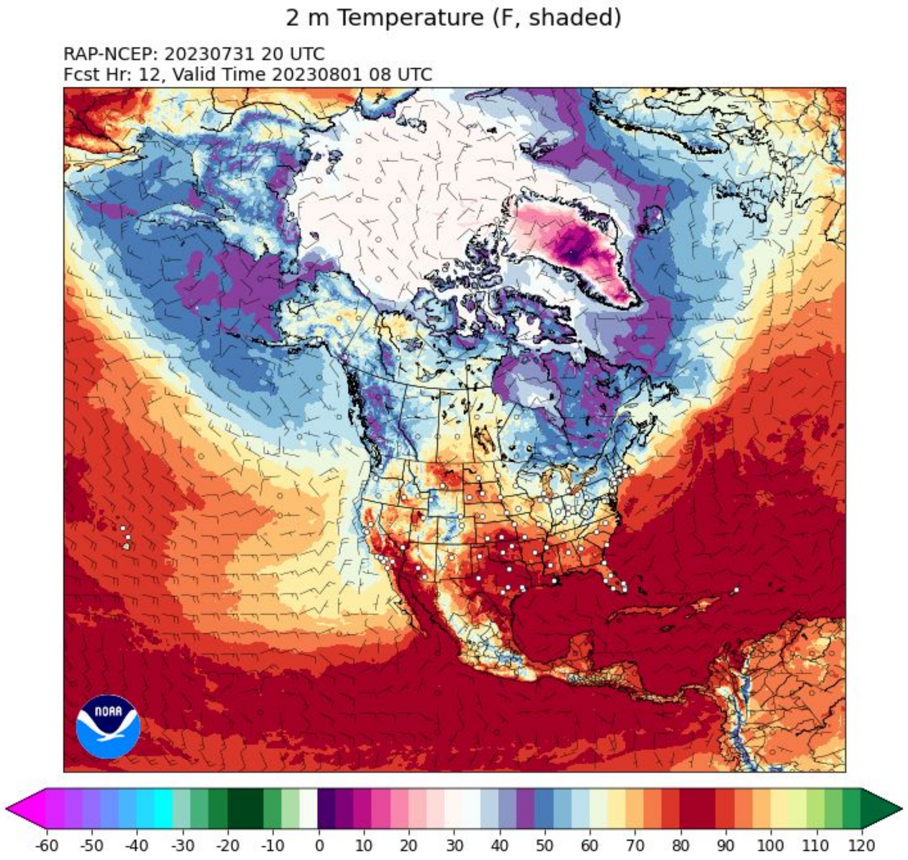
-> Profiles of atmospheric variables at each height (soundings):
- temperature
- dewpoint
- wind



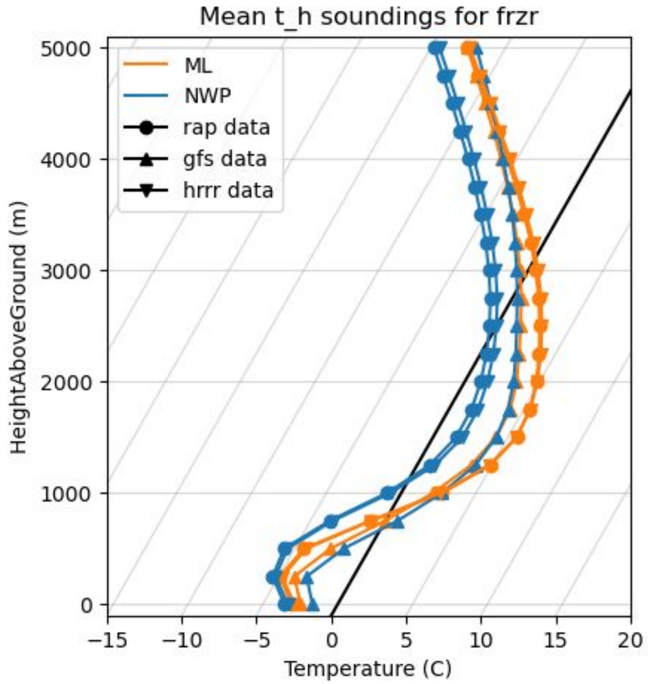
Where to get soundings?

Rapid Refresh (RAP) Model -> predictions for atmospheric variables

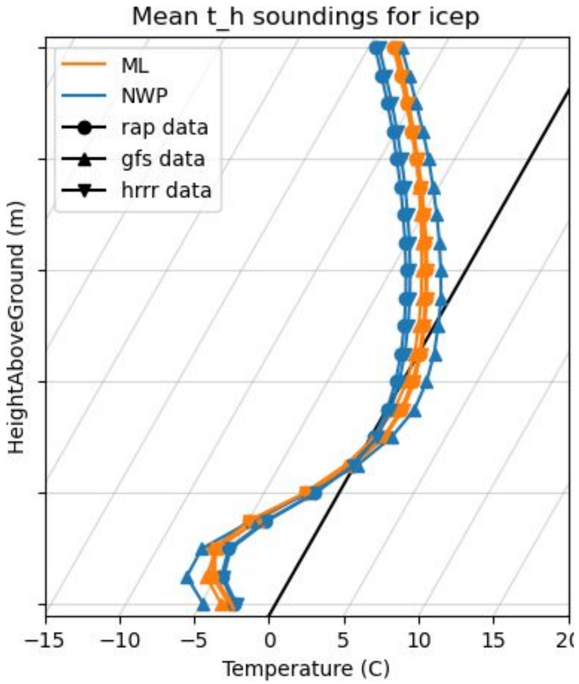




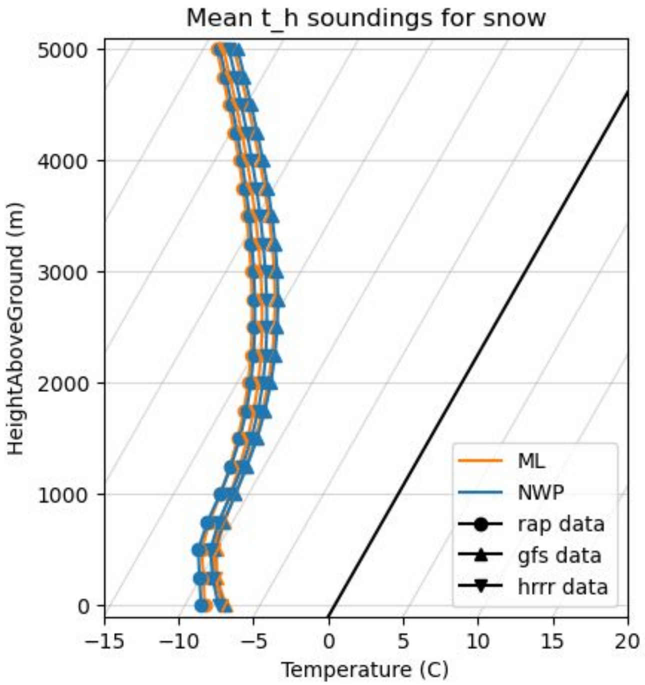
Evaluation: Composite Soundings



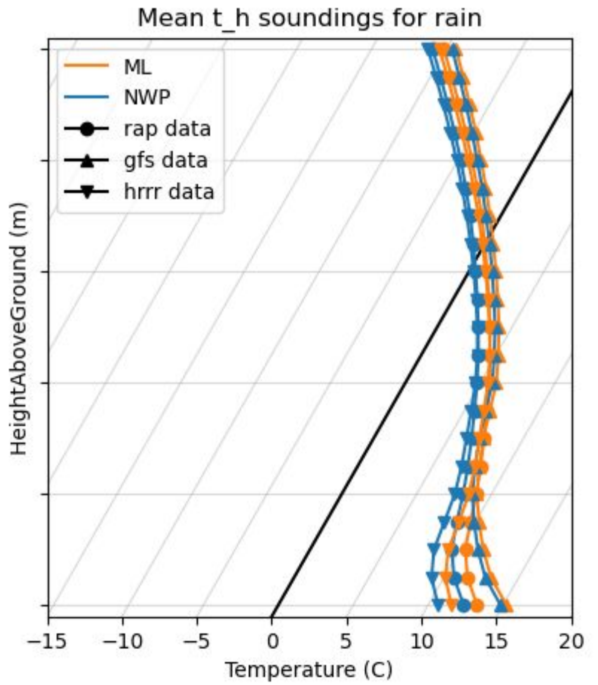
	rap	gfs	hrrr
ML frac_abv_0	0.97	0.96	0.96
NWP frac_abv_0	0.79	0.99	0.78
ML num_obs	2.1E+06	7.2E+04	3.7E+07
NWP num_obs	3.5E+06	6.1E+04	6.2E+07



	rap	gfs	hrrr
ML frac_abv_0	0.71	0.66	0.72
NWP frac_abv_0	0.89	0.94	0.89
ML num_obs	3.0E+06	1.3E+05	5.6E+07
NWP num_obs	8.7E+05	7.0E+04	1.6E+07



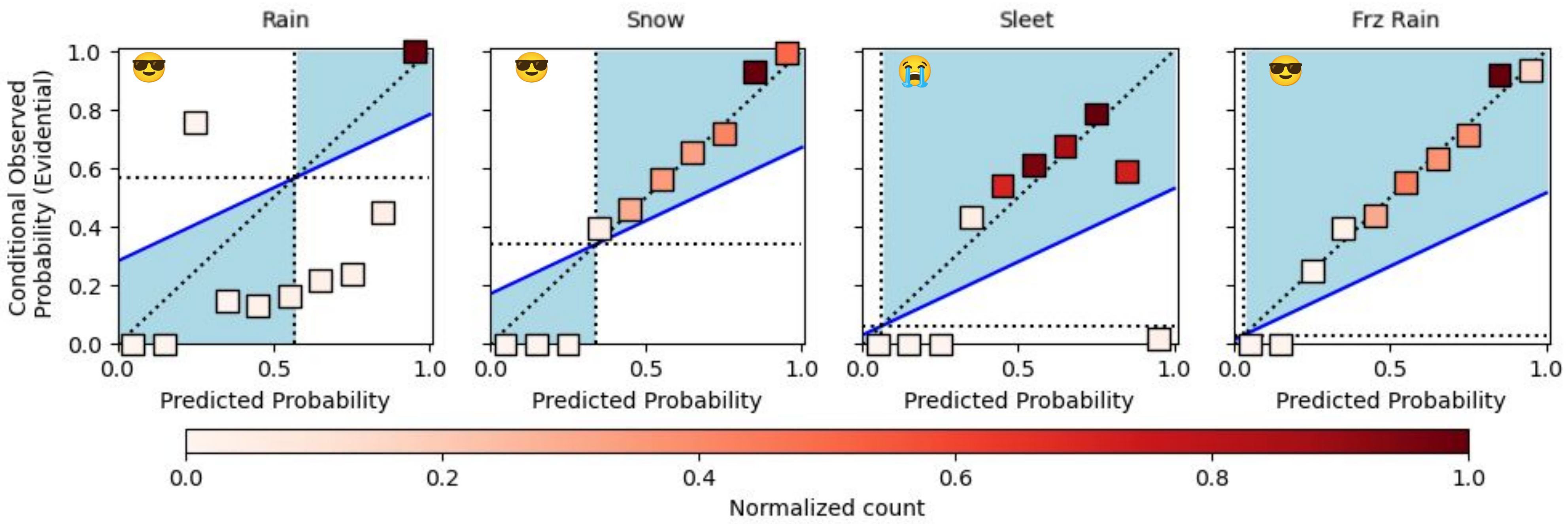
	rap	gfs	hrrr
ML frac_abv_0	0.10	0.11	0.11
NWP frac_abv_0	0.08	0.10	0.10
ML num_obs	5.2E+07	1.7E+06	6.7E+08
NWP num_obs	5.2E+07	1.7E+06	6.7E+08



	rap	gfs	hrrr
ML frac_abv_0	1.00	1.00	1.00
NWP frac_abv_0	1.00	1.00	1.00
ML num_obs	3.3E+07	1.7E+06	4.2E+08
NWP num_obs	3.6E+07	1.8E+06	4.6E+08

- Means are taken over 3TB of data
- Required significant engineering

Evaluation: Calibration

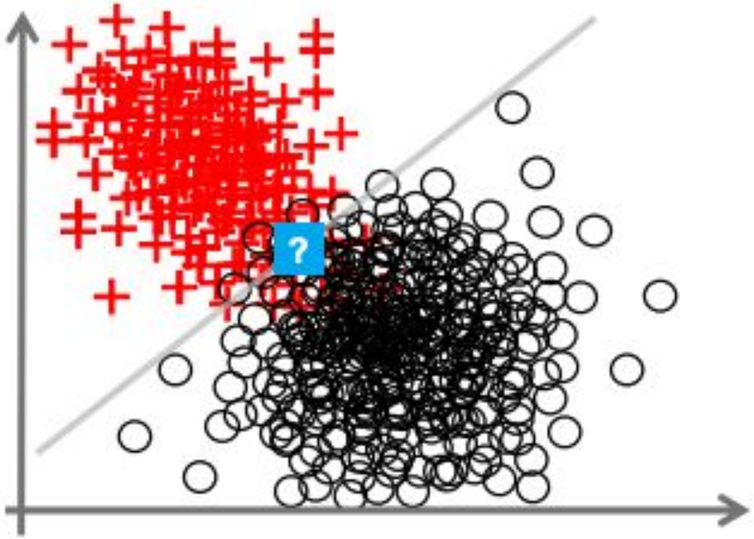


Ideal calibration curve: x=y line.

Why? ex. If model predicts label rain with probability p then true label should be rain p fraction of the time over examples the model predicts rain

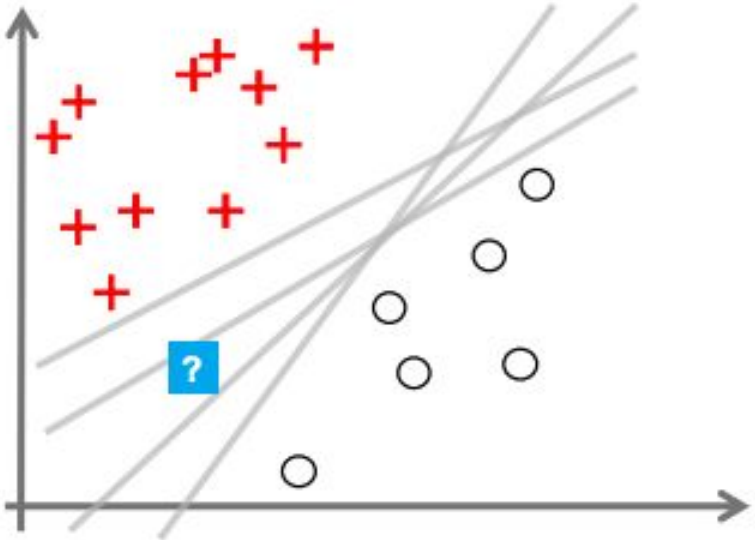
*figure by John Schreck of NCAR MILES group

Quick Aside: Uncertainty Quantification



Aleatoric Uncertainty

- irreducible
- inherent in the data

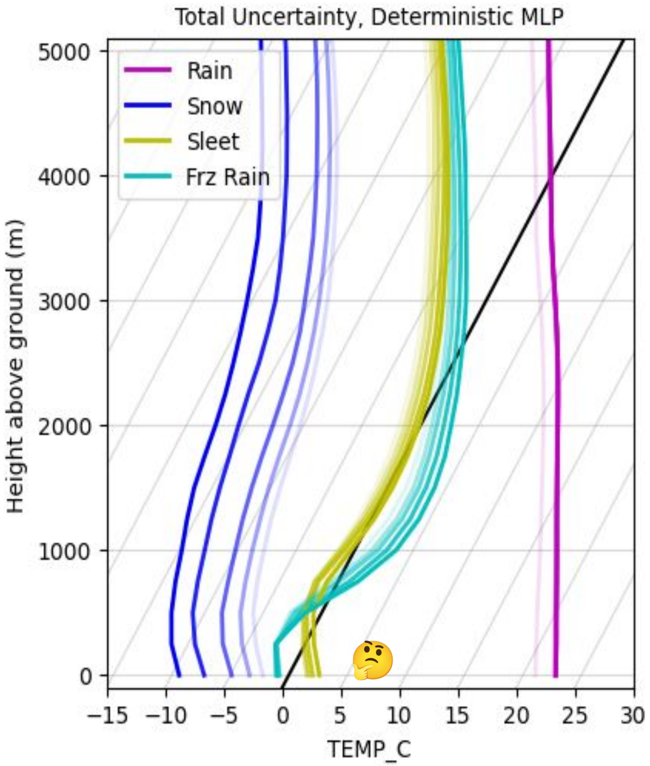


Epistemic Uncertainty

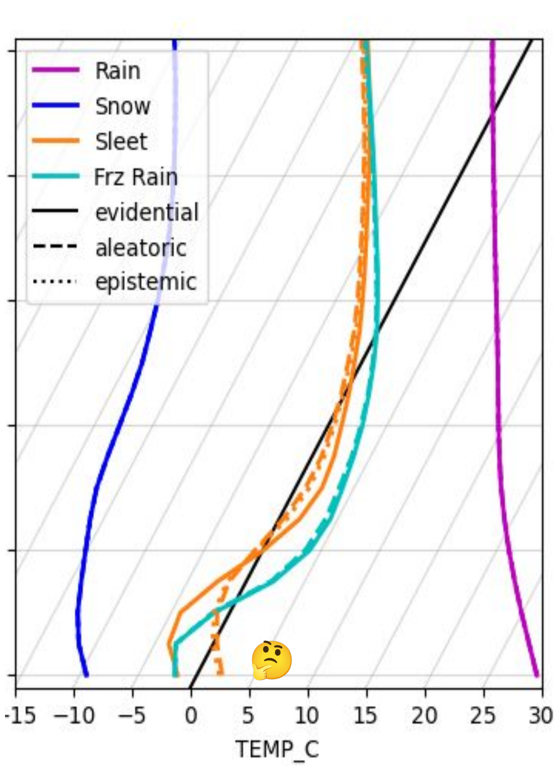
- reducible with more data, better modeling etc

Evidential Models can estimate uncertainty:
Sensoy, M., Kaplan, L., & Kandemir, M. (2018). Evidential deep learning to quantify classification uncertainty. Advances in neural information processing systems, 31.

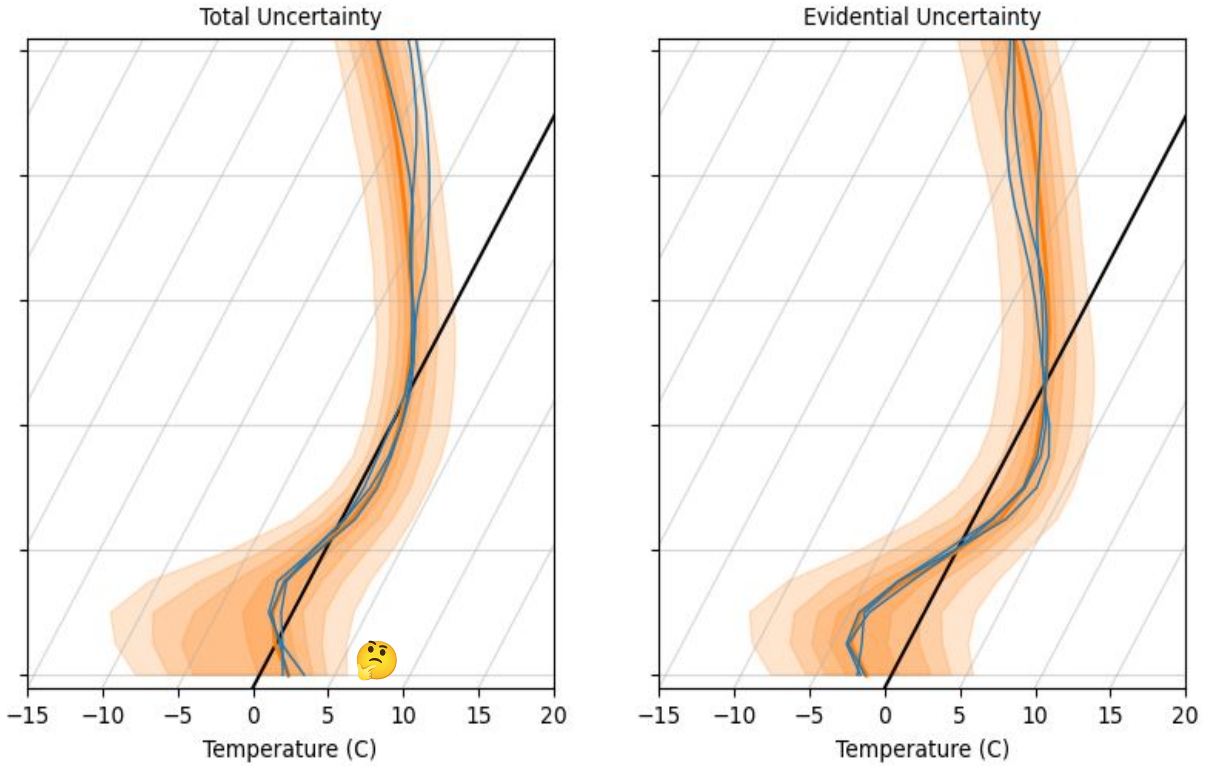
Evaluation: binned by uncertainty



MLP with Monte Carlo Dropout



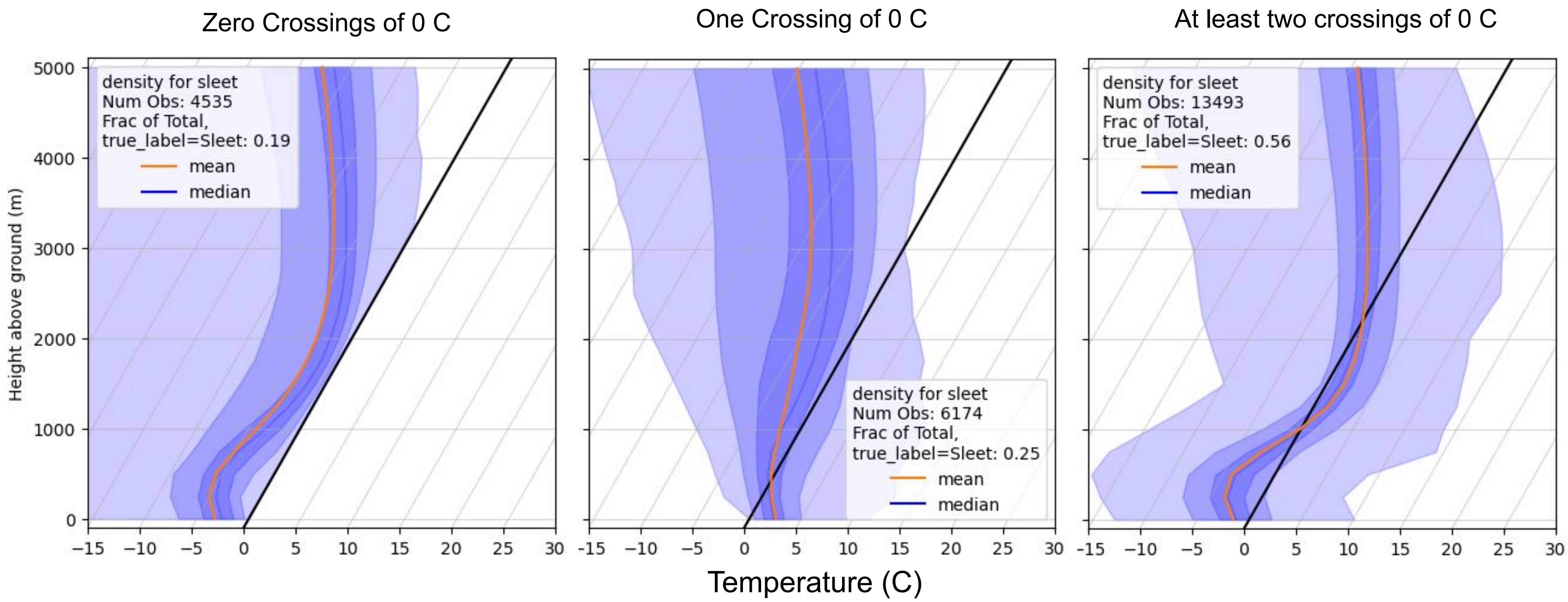
Evidential Model



Evidential Model, Sleet

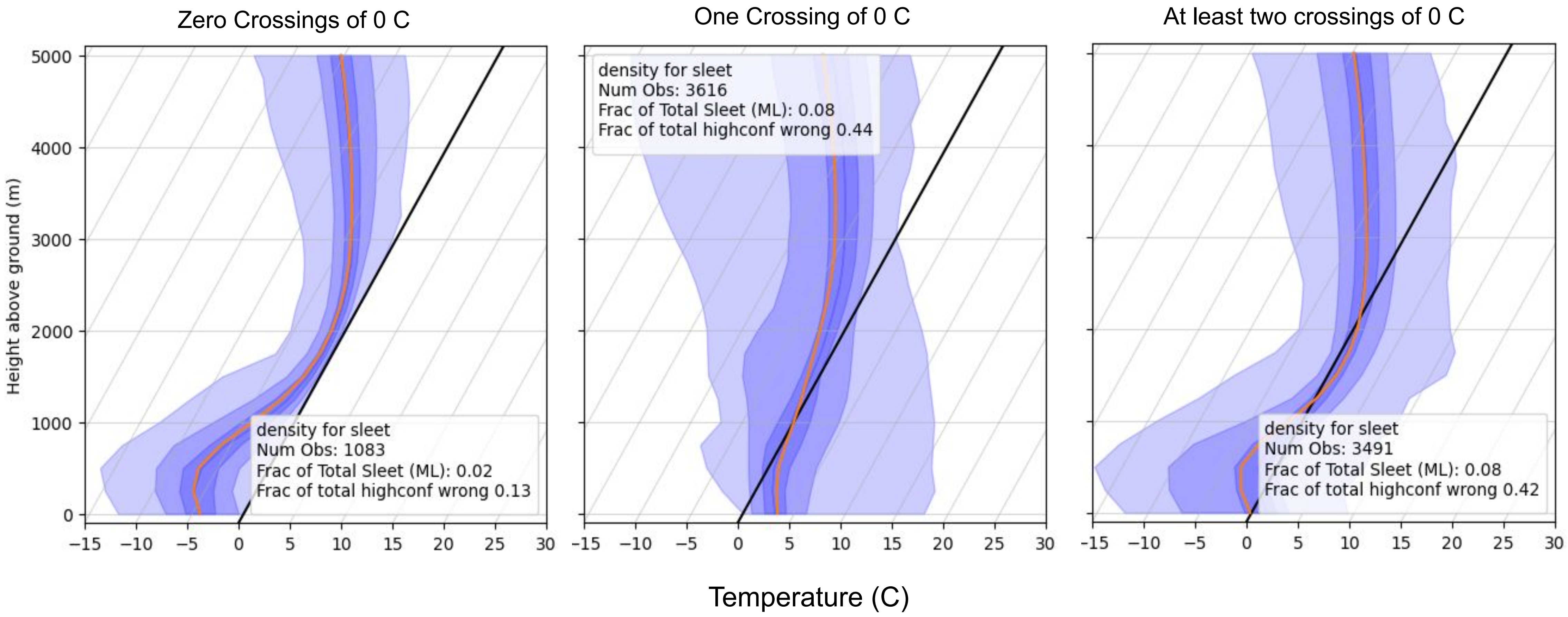
Root cause: Data Quality

- “ground truth” labels are from crowdsourced observations
- some quality control done, but not enough:



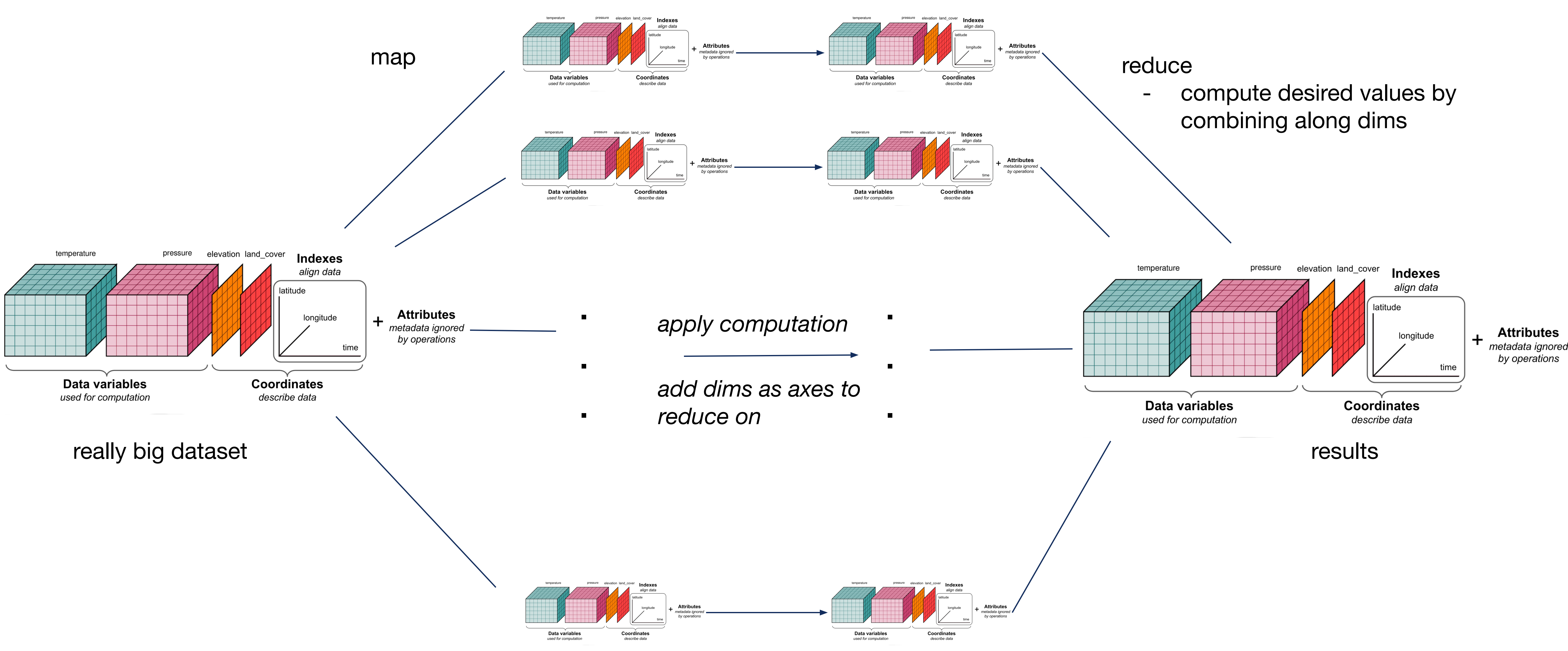
Root cause: Data Quality

- Soundings for high confidence and “wrong” ML predictions



End of evaluation section

Alternative approach to large Dask computations on large Xarray datasets



Pros/Cons

Pros

- less finicky than Dask which is very sensitive to chunking

- usually exists good approximations to expensive single-threaded computations e.g. histograms for quantile computation. Single threaded version: sort

Cons

- more user overhead
- not every function can be map-reduced
  - non parallelizable functions will be slow in dask also

Conclusion

Issues

- Which true labels for sleet are actually sleet?
- Evidential model has uncertainty blow-up

Future work

→ Further detailed investigation into convective precip. soundings

→ Use other NWPs for soundings

→ Improve loss function of evidential model

→ Hierarchical model to predict precip. type

→ Incorporate physics into model

Statistics

lines of code committed: 3396

file type	lines of code
.py	1484
.ipynb	1912

	Total	Per Business Day
CPU use	3400 core-hours	72 core-hours
RAM use	18594 gb-hours	395 gb-hours

Acknowledgements

Mentors:

- Charlie Becker
- David John Gagne
- Gabrielle Gantos
- John Schreck

Funding

- AI2ES
- NSF

dhamma.kimpara@colorado.edu