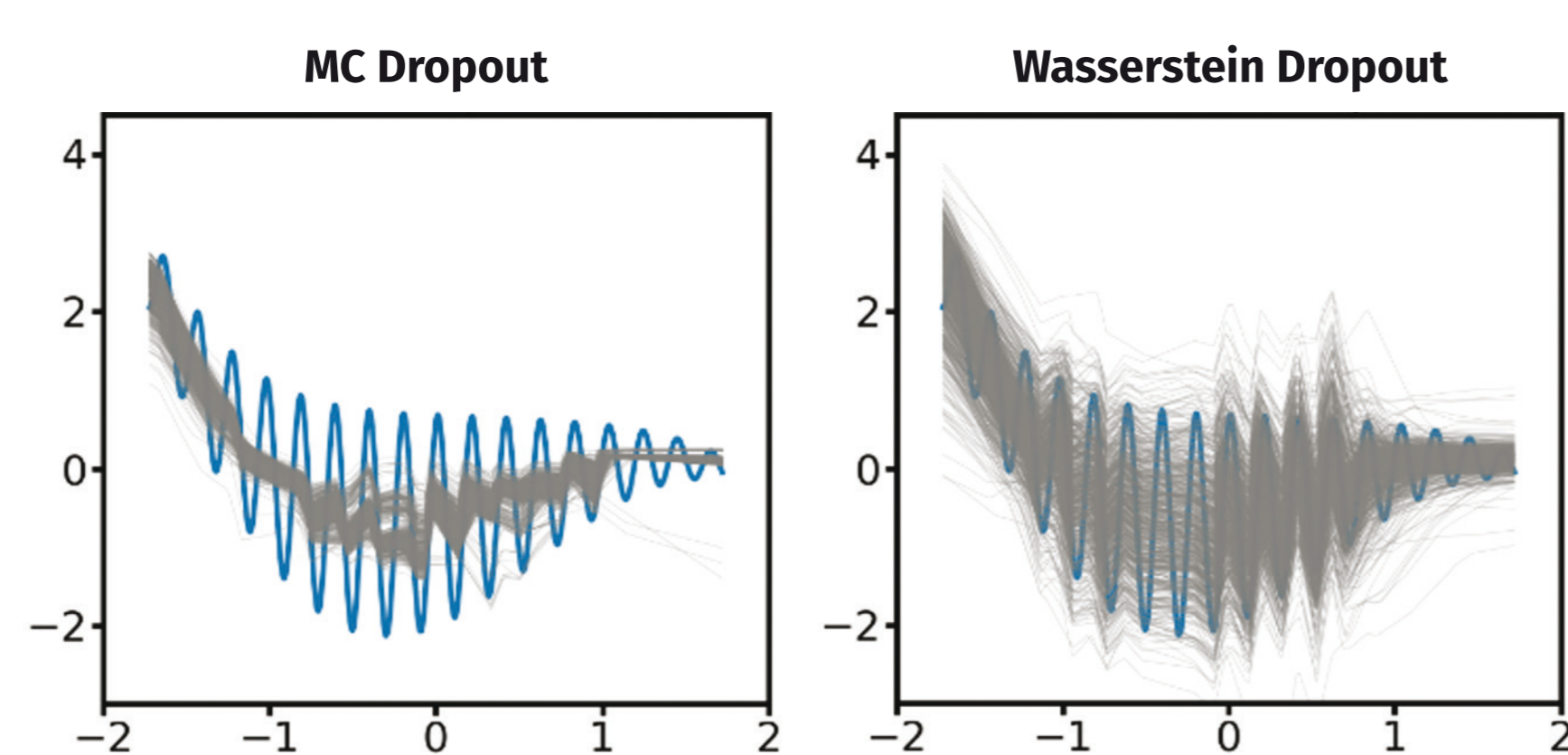


DNNs are fallible, especially when operated in open-world contexts where not all scenarios can be included in the training set. Uncertainty quantification attempts to determine whether a decision of a DNN can be trusted by estimating the probability or degree to which the decision is correct. For this, uncertainty estimates need to be local (i.e. per prediction) and well-calibrated (i.e. accurately estimated).

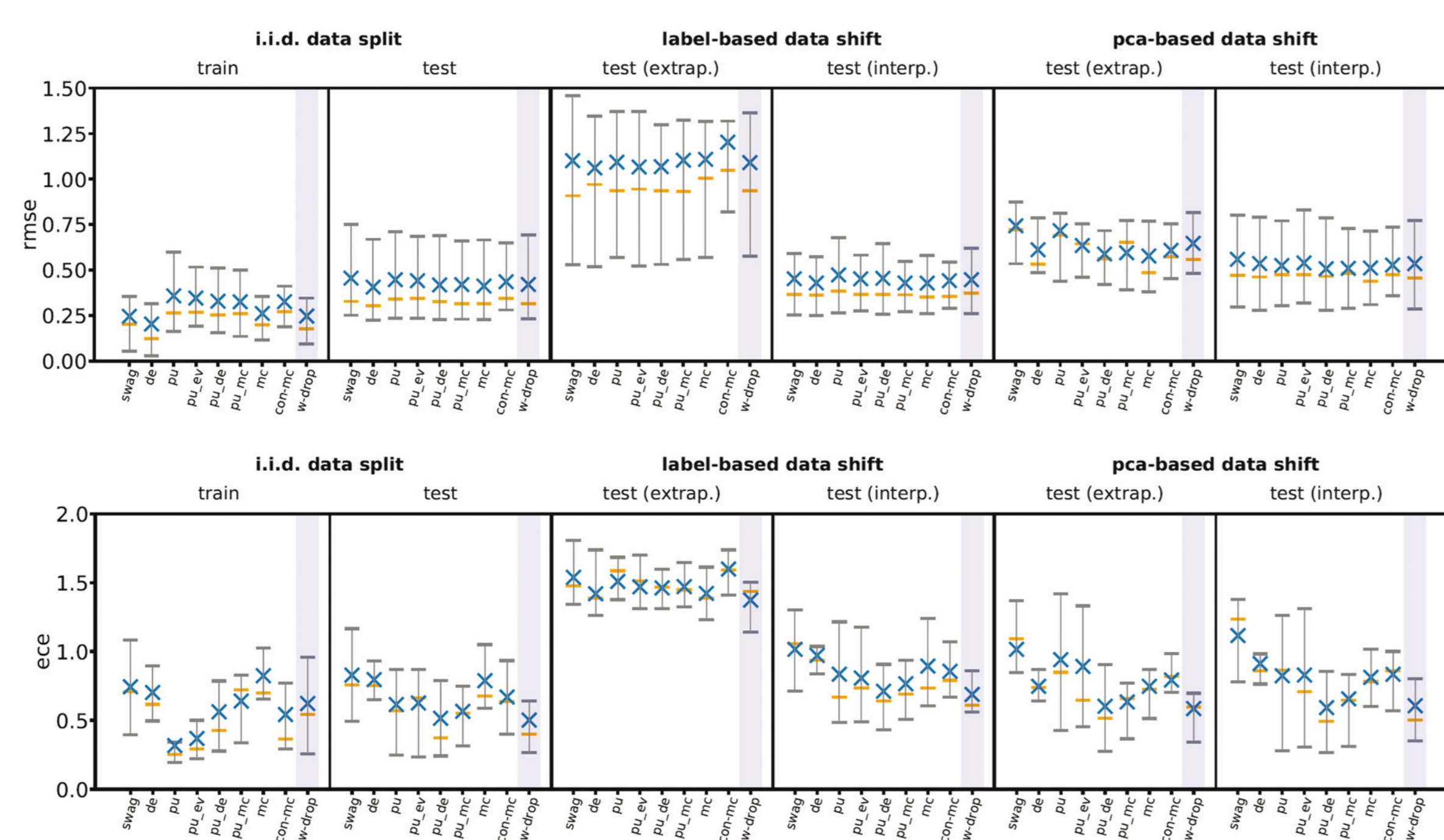
Conceptual Idea of Wasserstein Dropout

We extend dropout-based techniques, that are commonly used as regularizers and uncertainty estimators, to better reflect (varying) local uncertainty, as shown below. The idea is to directly control the (implicit) ensemble distribution generated by dropout and to tune its width to the data distribution. We measure the distance between both distributions using the Wasserstein-2 metric.



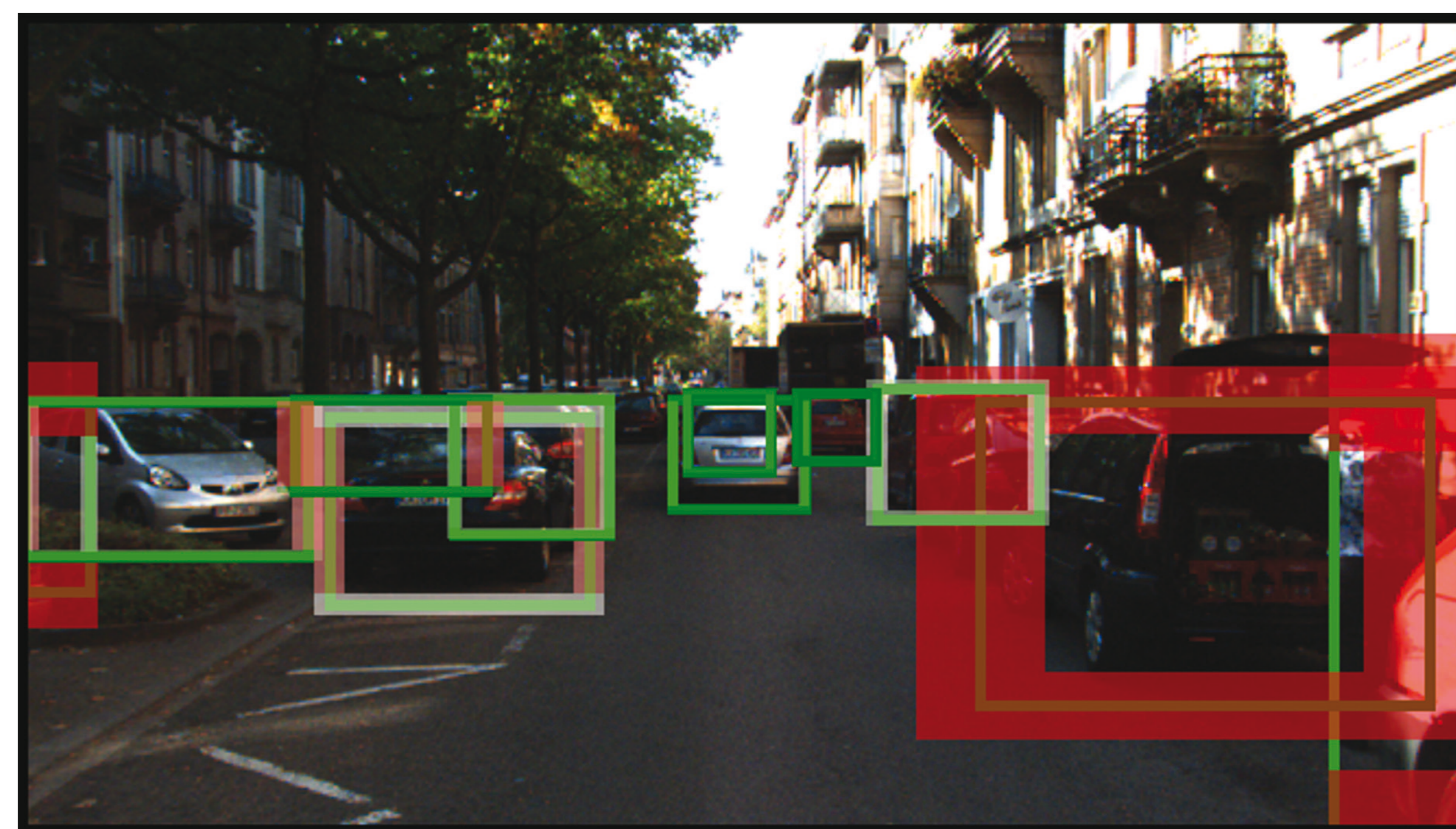
1D Case (UCI Regression)

Our approach, termed Wasserstein dropout, is adjusted for regression tasks and we evaluate it on multiple UCI benchmark datasets against various other estimation approaches. An overview is given below, where “cross” symbols denote means, and “whiskers” 25%/75% quantiles over the datasets. Regarding performance (RMSE) all methods are roughly on par on the test sets. The uncertainty quality (ECE), however, strongly differs, with our method producing competitive results on test datasets and under data shift. Please note that small 75%-quantiles indicate robustness across datasets.



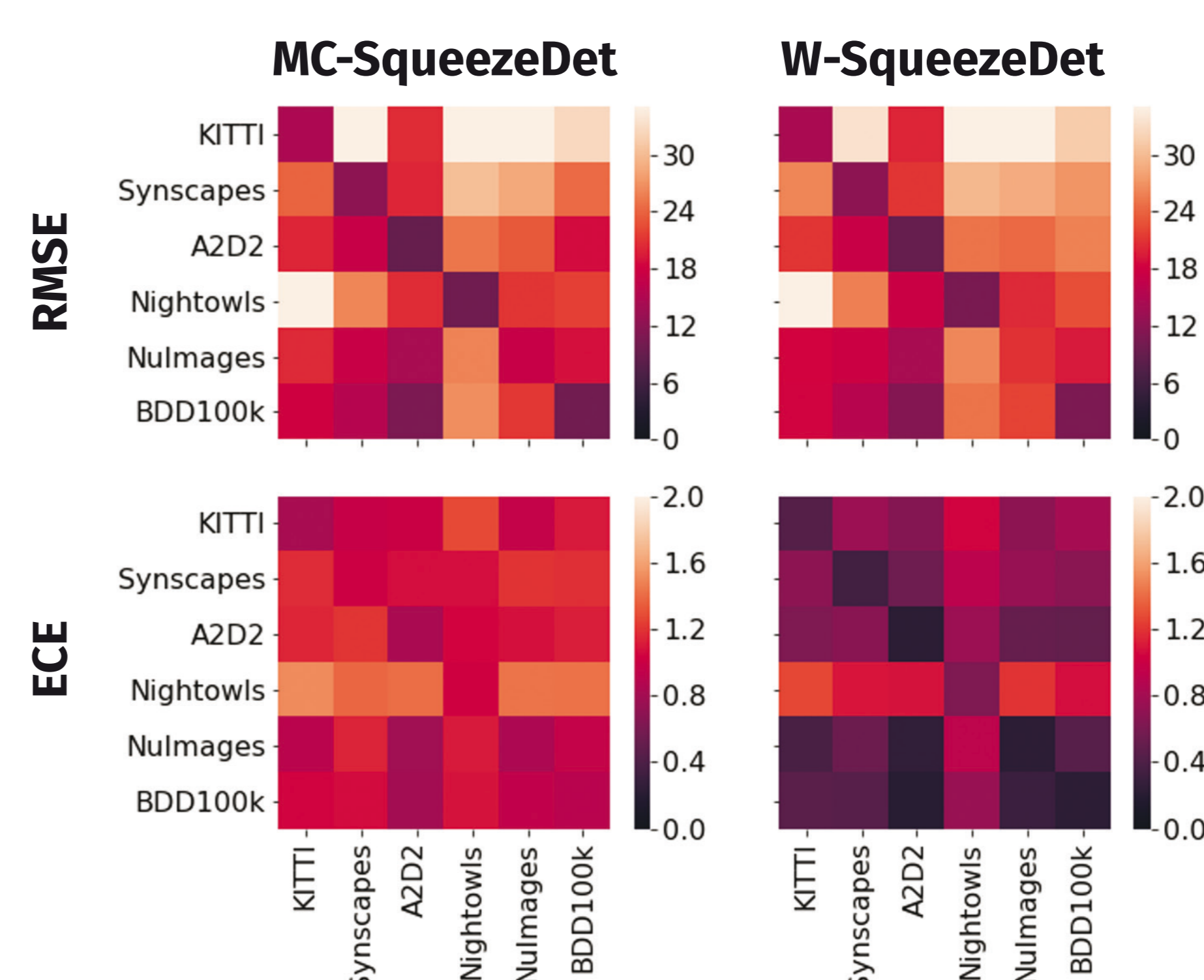
Safety Hypothesis:

DNN predictions may contain errors or inaccuracies. (Local) confidence information can be used to estimate the reliability of a given network output and thereby, potentially, safeguard against unsafe actions.



Object Detection

The prediction of bounding box coordinates, in parts, can be seen as a 4D regression problem to which we apply Wasserstein dropout. The figure above shows a scene from KITTI using SqueezeDet. Width and color of the boxes visualize uncertainty, which is particularly high for the rare case of a car with an open trunk (r.h.s.). Evaluations on further datasets show that Wasserstein dropout also works well for out-of-domain data.



Discussion & Outlook

Using our method on KI-A data and more complex DNNs (SSD, RetinaNet) revealed ~10% performance (mAP) degradation compared to the baseline. We suspect this is caused by the current algorithm to aggregate bounding box proposals and might be diminishable. Further work could also evaluate whether DNN errors are more amenable when hedged with uncertainty (compare MLRS20). Lastly, an extension to classification seems possible, e.g., to contribute to TP/FP distinctions.