Forecasts generated using protected time series change significantly from those that are generated using confidential data. While prior experiments have demonstrated severe degradations in forecast accuracy from a VAR model applied to differentially private time series, little is known about how data protection affects other forecasting models . We analyze the effects of several data protection methods (top and bottom coding, additive noise, differential privacy, and cluster-based swapping) on both simple and complex forecasting models.

We find that data protection generally degrades forecast accuracy regardless of forecast horizon.

Under differential privacy and additive noise, we find that exponential smoothing models have better accuracy than LGBM models for all forecast horizons.

However, under top and bottom coding, LGBM models have better accuracy than exponential smoothing models for longer forecast horizons.

We investigate the drivers of these results and offer guidance for practitioners on selecting a forecast model under various data protection approaches.