Protecting the privacy of time series data has the potential to expand forecasters’ access to otherwise unavailable datasets. However, forecasts generated using protected time series may change significantly from those that would have been generated using the confidential data. Prior experiments have demonstrated severe degradations in forecast accuracy using differentially private data with a VAR model, but it is not known whether using other privacy protection methods and forecasting models will result in similar degradations.

We analyze the effects of 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…

~~We relate the changes in forecasts resulting from data protection to the literature on judgmental forecasting, where significant downward adjustments in forecasts can produce large improvements in forecast accuracy.~~

* Just use forecast accuracy instead of forecast error.
* Use same terminology for improving or getting worse as was used in the fildes paper

In our preliminary results, we find that as expected, increasing the degree of top or bottom coding or random noise added to the data tends to reduce forecast accuracy. ~~In some cases, data protection can improve forecast accuracy for the majority of time series.~~ Surprisingly, we find that under top coding with exponential smoothing, the majority of forecasts are *upward* adjusted, and that the *downward* adjusted forecasts exhibit a larger relative increase in forecast error. Under bottom coding, *downward* adjusted forecasts tend to exhibit significant reductions in forecast error, especially under longer forecast horizons, where reductions in global error can range 25-30%. This mirrors past judgmental forecasting results where forecasts that are adjusted downward tend to be more accurate. Using LGBM with top coding, the *upward* adjusted forecasts exhibit increases in mean absolute error over short forecast horizons, and significant reductions in mean absolute error over long horizons. *Downward* adjusted forecasts see reductions in accuracy regardless of forecast horizon length. Protection through additive noise or differential privacy reduces forecast accuracy in the vast majority of cases.

The increases in relative forecast error under these methods are many magnitudes larger for *upward* adjusted forecasts than *downward* adjusted forecasts, regardless of the amount of noise added. We find that simpler models tend to generate more accurate forecasts than complex models when time series are protected using random noise and bottom coding. Our work represents a unique authority on the combinations of data protection methods and forecasting models that maximize forecast accuracy under privacy protection.

Under data protection, we recommend using simpler forecast models, such as exponential smoothing, which offer higher forecast accuracy.

Although simple forecasting models have worse accuracy on unprotected data, they outperform complicated models on protected data. This holds true across all horizons (1-20) and in addition, we find that accuracy improves at longer horizons but worsens at shorter horizons due to protection. We investigate the reasons in this paper and

The downward adjusted forecasts have better accuracy than those adjusted upward in short forecast horizons.