**The Effects of Privacy Protection on Forecast Accuracy**

Speaker: Cameron Bale, Ph.D. Student, LeBow College of Business, Drexel University

Coauthors: Matthew J. Schneider, Jinwook Lee

Forecasts generated using protected time series change significantly from those using the original time series. 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 privacy protection affects other forecasting models. We measure 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 degrades forecast accuracy the majority of the time regardless of forecast horizon. Surprisingly, when the time series are protected with differential privacy or additive noise, we find that exponential smoothing models have better accuracy than LGBM models for all forecast horizons. We investigate the reasons behind these results and offer guidance for practitioners in selecting a forecast model for privacy protected time series data.

**Introduction/Literature Review**

1. Effect of data protection on forecast accuracy is understudied
   1. Forecasting is popular in a variety of fields, such as consumer analytics, renewable energy and power industries, and census tracking. Privacy sensitive data such as UGC which have been shown to reduce forecast errors (Boone et al., 2019).
   2. However, legislation such as GDPR, California Consumer Privacy Act, California Privacy Rights Act, Colorado Privacy Act, Virginia Consumer Data Protection Act, and the Utah Consumer Privacy Act encourage organizations to protect time series data.
   3. Recent research found that the strongest form of privacy protection (differential privacy) degrades forecast accuracy of VAR models, but accuracy can improve when data owners collaborate and share their measurement data instead (Gonçalves et al., 2021).
   4. Other studies (Abowd et al., 2012; Nin & Torra, 2009) found that noise infusion had a small effect on auto-correlation coefficients and that the protected data was unbiased on average, while protection using *k*-anonymity produced large changes in the means, autocorrelation functions, and forecast accuracy of protected time series.
   5. In this paper, we care about how multiple forecasting models, from simple to complex, perform when trained on protected data.
2. Forecasts based on protected data are closely related to the judgmental forecasting (Petropoulos et al. 2021, sections 2.11.2 and 3.7.3)
   1. First, different forms of data protection produce multiple data points which have a different forecast. This differs from judgmental forecasting having multiple forecasts from a single data point.
   2. Second, similar to Fildes et al. (2009) we focus on cases where the forecast based on the protected data is improving or degrading forecast accuracy.
   3. Third, we require insights to minimize accuracy. Involves analyzing multiple data protection methods and forecasting models. We measure how data protection alters time series characteristics such as levels and trends.
3. How data protection changes time series characteristics and translate to changes in forecast accuracy
   1. Volatility of adjustment
      1. Judgmental adjustments can increase accuracy when based on reliable information, but information with low diagnosticity can damage forecast accuracy (Fildes et al., 2009, 2019).
      2. When accuracy did improve, increases in accuracy from adjustments to forecasts were greater for low volatility series which are easier to forecast (Fildes et al., 2009). We will examine the changes in forecast accuracy due to data protection with low diagnosticity (random noise).
      3. Adjusting for outliers, mainly those close to the forecast origin (similar to data protection), can increase forecast accuracy (Chen & Liu, 1993). Since top and bottom coding will reduce the effect of outliers on forecasts, we will examine how forecast accuracy changes for series with outliers.
4. How changes to forecasts are related to changes in forecast accuracy.
   1. Direction of adjustment.
      1. Both negative and positive adjustments increase accuracy (positive gives only marginal increase), overall accuracy increased by 10% (Davydenko & Fildes, 2013). Negative adjustments tended to reduce bias, while positive adjustments maintained bias magnitude or exacerbated it (Fildes et al., 2009).
      2. For judgmental forecasts, the direction of adjustment is chosen with the intent of improving the forecast – under data protection, direction is a result of the protection method, without regard to the effect on forecasts. We will measure the proportions of forecasts that are adjusted positively and negatively, the accuracy of the forecasts for each adjustment direction, and the resulting changes in forecast bias.
   2. Size of adjustment.
      1. (Fildes et al., 2009) found that the size of adjustments was positively associated with the size of accuracy improvements. One explanation given is that an adjuster who is more confident will make a larger adjustment, which improves accuracy when based on reliable information.
      2. Under data protection, the adjustment size will be determined by the forecasting models’ responses to the data protection, which are likely related to the strength of protection.
      3. We will measure the strength of protection by the magnitude of the adjustment (P\_t-A\_t), how this relates to the size of adjustments in the forecasts based on protected data, and whether adjustment size is related to the change in forecast accuracy from data protection. We need to be careful with large positive adjustments versus large negative adjustments.
5. Extant studies on data privacy and forecasting.
   1. (Goncalves et al., 2021) found that data owners have a monetary incentive to share their data, but may be discouraged from doing so due to privacy concerns over sharing data with a central party. Transfer learning (Wellens et al., 2021) is designed to train models on sensitive data and transferred themselves, which can overcome privacy barriers or regulations. Our work would help answer how forecast accuracy would be affected if the data owners applied data protection methods prior to sharing their data using single protected dataset, not shared models which are computationally expensive.
   2. Differential privacy is the most popular data protection method applied to time series, see (Imtiaz et al., 2020; Liyue Fan & Li Xiong, 2014) for examples. An interesting result from (Imtiaz et al., 2020) is that differentially privacy data did not always produce worse forecast accuracy when forecasting individuals’ health data using a recurrent neural network. The reason is adding random Gaussian noise to time series is a technique to prevent overfitting when forecasting with neural networks (Hewamalage et al., 2021).
   3. Data integrity attacks (Luo et al., 2018) add random noise to data and have also been found to degrade accuracy of multiple linear regression, artificial neural network, support vector regression, and fuzzy interaction regression models. Support vector regression was most robust and performed the best under the data integrity attacks.
   4. We will explore whether data protection using differential privacy improve forecast accuracy for neural networks and simpler models alike and underlying reasons *why* the forecast accuracy changed.
6. Data protection literature review
   1. Additive/Multiplicative noise is similar to differential privacy in that random noise is infused into the data, but there are no theoretical privacy guarantees. (Luo et al., 2018) use multiplicative random noise to simulate a data integrity attack in load forecasting time series. (Abowd et al., 2012) protect individual series in QWI data using a single multiplicative noise factor that is applied to all time periods.
   2. *K*-anonymity is a definition where every record (or time series) in a dataset is identical to at least *k*-1 other records (Sweeney, 2002).   
      (Nin & Torra, 2009) evaluate the change in forecast accuracies of simple exponential smoothing, double exponential smoothing, linear regression, multiple linear regression, and polynomial regression applied to *k*-anonymized data and found significant reductions in forecast accuracy across all five models even for *k*=2, they do not provide a comparison between models or additional data protection methods.
   3. (Crimi & Eddy, 2014) study the effect of top coding the Census’ Public Use Microdata Samples on analyses of interest. They find that the sample correlation between two variables is shrunk towards zero when one or both of the variables are top coded. We seek to explore whether these changing correlational structures affect multivariate forecasting model performance.
   4. New shuffling technique (ours) based on time series similarities? (we could use from other paper)
7. Our contributions (to be decided after results)
   1. An empirical comparison of forecasting models’ accuracies based on protected data. We seek to discover whether simple or complex models forecast more accurately after data protection.
   2. We apply forecasting models to the M3 monthly micro data. In addition, we use time series characteristics, such as spectral entropy, to explain the changes in model performance post-data protection.

**Empirical Application (or Data/Model/Results sections)**

1. **Data** – M3 monthly Micro
   1. Describe data and the original competition. 474 time series, min/max
   2. Complex models perform well on unprotected data (Cite literature here)
   3. Models that explicitly model trend and seasonality performed the best (Kang et al., 2017)
   4. Measure levels/trends/seasonalities on the unprotected data. Measure all time series characteristics, AR(1) parameters, frequency (monthly), etc.
   5. \*We could benefit from a better data set for managerial implications\*
2. **Methods** – If we are using k-means shuffling, we can have this here
   1. Table including data protection methods - - Differential privacy, top/bottom coding, additive/multiplicative noise
   2. Specify how time series characteristics are used to decide which time series are most similar
   3. Include a proof or two showing why it should improve accuracy
3. **Models** 
   1. Selection of Forecasting Models
      1. Popular models found in R, Python and/or serving as benchmarks in forecasting competitions. These models show how forecast accuracy would change for ‘non-expert’ users.
      2. SES, DES, TES, ARIMA, LGBM (univariate), VAR, LGBM (multivariate), RNN (LSTM cells)
      3. Prophet (Facebook), Greykite (LinkedIn), NeuralProphet
      4. Discuss briefly about tuning parameters selected/etc. – this is not the focus of the paper.
4. **Analysis Framework**
   1. Forecasting Process
      1. Fixed origin forecasting with two horizon lengths: 1-step and 18-step (18 step was used for monthly data in M3 (Makridakis & Hibon, 2000))
      2. Describe process, train, test, validation sets
         1. Obtain original series **,** Create protected series
         2. Extract time series characteristics from and
         3. Train each forecasting model on and
         4. Generate forecasts for and
   2. Measure forecast accuracies as categorized by the intro/lit review
      1. Overall – for all time series
         1. MAE for directly comparing model accuracy.
         2. AvgRelMAE (scale independent, robust to outliers, easily interpretable) for comparing accuracy change after data protection: , where , and where denotes protected, denotes original, is the number of adjusted forecasts for series and is the number of series. AvgRelMAE < 1 (> 1) indicates increased (damaged) accuracy on average. Average percentage improvement in MAE of forecasts is (1 – AvgRelMAE) 100. 5% trim on AvgRelMAE is recommended (Davydenko & Fildes, 2013).
      2. Explain why accuracies are different using time series features and visualize using instance space (PCA applied to extracted features) (Hewamalage et al., 2022; Kang et al., 2017).
         1. Examine distributions of time series characteristics (e.g., seasonality, trend) before and after data protection to explain *why* models had increases/damages in accuracy.
         2. Spectral entropy (forecastability/diagnosticity).
            1. Relate to 3.1.1 and 3.1.2.
         3. Coefficient of variation (directly measure volatility) (Fildes et al., 2009).
            1. Relate to 3.1.1 and 3.1.2.
         4. Strength of trend and seasonality.
         5. First order autocorrelation.
         6. Optimal box-cox transformation parameter.
         7. Outliers – how accuracy changes for series with outliers and how that depends on distance of outliers to forecast origin
            1. Relate to 3.1.3.
      3. Direction - Proportion of positive/negative adjusted forecasts. Change in accuracy (AvgRelMAE) for positive and negative adjusted forecasts. Bias (mean of forecast errors) for positive and negative adjusted forecasts.
         1. Relate to 4.1.2 and table 3 below.
      4. Magnitude
         1. Measure magnitude of data adjustment ||
         2. Measure size of adjustment in forecasts: Size= 100 ×|Protected-Original|\/Original (Fildes et al., 2009).
         3. Measure relationship between size of adjustment and change in forecast accuracy.
         4. Look at this overall and split into positive adjustments and negative adjustments.
   3. Complexity – relationship between complexity and forecast accuracy before and after data protection. Measure complexity using Kolmogorov complexity as described by (Hewamalage et al., 2022)
      1. Train Model 🡪 Save Model (coefficients, objective functions, etc.) 🡪 Gzip Compressed Model 🡪 Approximate Kolmogorov Complexity using size of the compressed model on disk.
      2. Regress Accuracy ~ Complexity. To make conclusive statements on how model complexity affects accuracy using protected data. Cite prior literature for the unprotected data (complex models forecast more accurately!)

**Results**

**We need tables that forecasting readers want to see here. You can have the detailed ones as well ,but think hard about overall metrics to display based on the intro/lit reviews. These are usually found in other research papers but you need to be creative too in order to put ones in the results subsections above.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Original | |  | |  | |  | |  | |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| SES |  |  |  |  |  |  |  |  |  |  |  |  |
| DES |  |  |  |  |  |  |  |  |  |  |  |  |
| TES |  |  |  |  |  |  |  |  |  |  |  |  |
| ARIMA |  |  |  |  |  |  |  |  |  |  |  |  |
| Prophet |  |  |  |  |  |  |  |  |  |  |  |  |
| Greykite |  |  |  |  |  |  |  |  |  |  |  |  |
| LGBM (univariate) |  |  |  |  |  |  |  |  |  |  |  |  |
| VAR |  |  |  |  |  |  |  |  |  |  |  |  |
| LGBM (multivariate) |  |  |  |  |  |  |  |  |  |  |  |  |
| NeuralProphet |  |  |  |  |  |  |  |  |  |  |  |  |
| RNN (LSTM cells) |  |  |  |  |  |  |  |  |  |  |  |  |

Adapt the following table for our purposes (Fildes et al., 2009).

A screenshot of a computer

Description automatically generated with medium confidence

Two-sided binomial test for the probability of an adjustment that increases accuracy – use similar table as below (Davydenko & Fildes, 2013).

A picture containing Word

Description automatically generated

**Conclusions (incomplete)**

1. General Discussion
   1. Which models performed best on original data.
      1. Why.
   2. Which models performed best on protected data.
      1. Why.
2. Managerial Implications
   1. Limit analysts to simple models when forecasting using protected data.
3. Future Work
   1. Ways to improve forecast accuracy under data protection.
      1. Make adjustments based on methods from judgmental forecasting literature (Fildes et al., 2009)

Like forecasting support systems that actively evaluate and filter information before presenting it (Fildes et al., 2019), design (or implement) methods to filter random noise from the series prior to forecasting.

**References**

IAPP, 2020. US State Privacy Legislation Tracker. <https://iapp.org/resources/article/us-state-privacy-legislation-tracker/>.

Abowd, J. M., Gittings, K., McKinney, K. L., Stephens, B. E., Vilhuber, L., & Woodcock, S. (2012). *Dynamically consistent noise infusion and partially synthetic data as confidentiality protection measures for related time-series*. 41.

Boone, T., Ganeshan, R., Jain, A., & Sanders, N. R. (2019). Forecasting sales in the supply chain: Consumer analytics in the big data era. *International Journal of Forecasting*, *35*(1), 170–180. https://doi.org/10.1016/j.ijforecast.2018.09.003

Chen, C., & Liu, L.-M. (1993). Forecasting time series with outliers. *Journal of Forecasting*, *12*(1), 13–35. https://doi.org/10.1002/for.3980120103

Crimi, N., & Eddy, W. (2014). Top-Coding and Public Use Microdata Samples from the U.S. Census Bureau. *Journal of Privacy and Confidentiality*, *6*(2). https://doi.org/10.29012/jpc.v6i2.639

Davydenko, A., & Fildes, R. (2013). Measuring forecasting accuracy: The case of judgmental adjustments to SKU-level demand forecasts. *International Journal of Forecasting*, *29*(3), 510–522. https://doi.org/10.1016/j.ijforecast.2012.09.002

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: An empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, *25*(1), 3–23. https://doi.org/10.1016/j.ijforecast.2008.11.010

Fildes, R., Goodwin, P., & Önkal, D. (2019). Use and misuse of information in supply chain forecasting of promotion effects. *International Journal of Forecasting*, *35*(1), 144–156. https://doi.org/10.1016/j.ijforecast.2017.12.006

Gonçalves, C., Bessa, R. J., & Pinson, P. (2021). A critical overview of privacy-preserving approaches for collaborative forecasting. *International Journal of Forecasting*, *37*(1), 322–342. https://doi.org/10.1016/j.ijforecast.2020.06.003

Goncalves, C., Pinson, P., & Bessa, R. J. (2021). Towards Data Markets in Renewable Energy Forecasting. *IEEE Transactions on Sustainable Energy*, *12*(1), 533–542. https://doi.org/10.1109/TSTE.2020.3009615

Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent Neural Networks for Time Series Forecasting: Current status and future directions. *International Journal of Forecasting*, *37*(1), 388–427. https://doi.org/10.1016/j.ijforecast.2020.06.008

Hewamalage, H., Bergmeir, C., & Bandara, K. (2022). Global models for time series forecasting: A Simulation study. *Pattern Recognition*, *124*, 108441. https://doi.org/10.1016/j.patcog.2021.108441

Imtiaz, S., Horchidan, S.-F., Abbas, Z., Arsalan, M., Chaudhry, H. N., & Vlassov, V. (2020). Privacy Preserving Time-Series Forecasting of User Health Data Streams. *2020 IEEE International Conference on Big Data (Big Data)*, 3428–3437. https://doi.org/10.1109/BigData50022.2020.9378186

Kang, Y., Hyndman, R. J., & Smith-Miles, K. (2017). Visualising forecasting algorithm performance using time series instance spaces. *International Journal of Forecasting*, *33*(2), 345–358. https://doi.org/10.1016/j.ijforecast.2016.09.004

Liyue Fan & Li Xiong. (2014). An Adaptive Approach to Real-Time Aggregate Monitoring With Differential Privacy. *IEEE Transactions on Knowledge and Data Engineering*, *26*(9), 2094–2106. https://doi.org/10.1109/TKDE.2013.96

Luo, J., Hong, T., & Fang, S.-C. (2018). Benchmarking robustness of load forecasting models under data integrity attacks. *International Journal of Forecasting*, *34*(1), 89–104. https://doi.org/10.1016/j.ijforecast.2017.08.004

Makridakis, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions and implications. *International Journal of Forecasting*, *16*(4), 451–476. https://doi.org/10.1016/S0169-2070(00)00057-1

Nin, J., & Torra, V. (2009). Towards the evaluation of time series protection methods. *Information Sciences*, *179*(11), 1663–1677. https://doi.org/10.1016/j.ins.2009.01.024

Sweeney, L. (2002). k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *10*(05), 557–570. https://doi.org/10.1142/S0218488502001648

Wellens, A. P., Udenio, M., & Boute, R. N. (2021). Transfer learning for hierarchical forecasting: Reducing computational efforts of M5 winning methods. *International Journal of Forecasting*, S0169207021001606. https://doi.org/10.1016/j.ijforecast.2021.09.011