Research Article

PUBLIC FINANCE JOURNAL

Modeling Approach Matters, But Not as Much as Preprocessing:
Comparison of Machine Learning and Traditional Revenue Forecasting Techniques

Sarah E. Larson[1, c], Michael Overton[2]

Revenue forecasting accuracy is critical to governmental operations. This paper addresses the question: What is the best technique for forecasting sales tax revenue? Prior studies in this area have focused on the differences between machine learning techniques and traditional approaches and neglected to consider how differences in pre-processing steps for the data before the forecasting model is applied are important. Here, we show that machine learning techniques do not always provide increased forecasting accuracy. Instead, the modeling choices matter, but less than the prior literature and practice suggested. Rather, pre-processing makes the most significant difference in forecasting accuracy, and forecasters need to understand the unique characteristics of time series data to improve forecasting performance. The immediate implications of these findings are that the focus of practitioners of in sales tax revenue forecasting should shift from prioritizing model choice towards data pre-processing.

Keywords: Machine Learning, Revenue Forecasting, Sales Taxation

What is the best technique to forecast revenue? State and local governments have grappled with this question for decades (Grizzle & Klay, 1994; Rodgers & Joyce, 1996; Rubin, Mantell, & Pagano, 1999). Differences in forecasting practices exist across countries (Buettner & Kauder, 2010), and overall forecasting accuracy can depend on political and organizational influences (Bretschneider & Gorr, 1992; Bretschneider et al., 1989).

The prior literature has also attempted to identify the accuracy of various methods (Boyd & Dadayan, 2014; Grizzle & Klay, 1994), such as regression (Wong, 1995) and time series models (Frank, 1990). Mikesell (2011) even advocated that "having the experience of 'old hands,' who have seen almost everything play out at least once before, can be crucial to getting

[1]Department of Political Science, Miami University. https://orcid.org/0000-0002-9644-2019. [2]Department of Politics and Philosophy, University of Idaho.. https://orcid.org/0000-0002-1800-9164. [c]Corresponding Author: larsonse@miamioh.edu.

those impacts correctly in the forecast" (p. 569), highlighting the importance of the human component and individual forecaster knowledge in the process. Yet, the accuracy of different methodological approaches depends on the unit of analysis, with damped trend analysis and exponential smoothing performing the best when forecasting monthly and quarterly data and naive approaches when forecasting switched to annual data (Williams & Kavanagh, 2016).

Recently, researchers have explored the accuracy of machine learning techniques as new revenue forecasting tools for state and local governments. Buxton et al. (2019) found that the two deep learning approaches, simple Multi-Layer Perception and global Multi-Layer Perception models, outperform the traditional moving average approach to forecasting for sales tax data within Illinois. Their study also broke out sales tax revenue by type, dividing the information into ten different categories of sales tax revenue within the state. In contrast, Chung, Williams, and Do (2022) focus on multiple types of local government revenue sources and the accuracy of machine learning forecasting techniques when compared to methods traditionally used by forecasters at the local level. They found that a machine learning method, k-nearest neighbor, performed the best at forecasting property taxes. In contrast, they found that a traditional forecasting approach dampened trend exponential smoothing, providing the best accuracy for forecasting sales taxes. In summation, the ability of the machine and deep learning methods to forecast with greater accuracy than traditional forecasting techniques remains inconclusive.

This paper addresses the question: What is the best technique for forecasting sales tax revenue? Like Buxton et al. (2019) and Chung et al. (2022), we focus on the differences in performance accuracy between traditional, non-machine learning techniques and machine learning techniques. However, methods are only as good as the quality of data used to analyze them (Cole, 1969). In this vein, we include an exploration of pre-processing steps–tasks taken to improve the quality of data prior to any analysis–across different forecasting time periods using monthly, quarterly, and annual sales tax data. Understanding that the ideal number and type of pre-processing steps might vary across forecasting time periods, we look at each time period separately. Therefore, the research question can be expanded to two questions. First, what is the best technique for forecasting sales tax revenue in both the number and type of pre-processing steps to the data and the forecasting method? Second, how do those findings change when the data are forecasted monthly versus quarterly versus annually?

The modeling and pre-processing techniques are defined in Table 1. The modeling (i.e., forecasting) techniques selected include three classical methods, three machine learning methods, and four benchmark methods to help us understand the importance of more sophisticated classical and machine learning style forecasting methods. These are listed in order in the table. The three classical or traditional methods are autoregressive integrated moving average (ARIMA), dampened trend exponential smoothing (DT ETS), and linear trend model. These methods were selected to provide approaches from both the time-series analysis or causal-like approaches found within traditional forecasting and have been common forecasting techniques for decades. The three machine learning approaches are K-nearest neighbor (KKNN), neural network autoregressive (NNAR), and extreme gradient boosting (XGBOOST). The three machine learning approaches use different algorithms through a training process to find underlying patterns and relationships in the data. Finally, the four benchmarking methods are used to create a series of baseline forecasts and include drift, naïve, seasonal naïve, and average methods.

The pre-processing actions included in this study are either general data pre-processing steps or those specifically required for time series data. General pre-processing steps include

Table 1. Model and Preprocessing Description

Type	Abbreviation	Long Name	Description
Preprocessing			
General	IHS	Inverse Hyperbolic Sine	Natural logarithm of a case plus the square root of a case-squared plus one.
General	Log	Natural Log	Natural logarithmic transformation of time series data.
Time Series	SA	Seasonally Adjusted	Average seasonal detrended time series obtained using multiplicative classical decomposition.
Time Series	Detrend	Detrend	Time series with the trend removed using multiplicative classical decomposition.
Model			
Traditional	ARIMA	Autoregressive Integrated Moving Average	Forecast using an automated process to determine the need and values for time series data differencing, autoregression, and moving averages.
Traditional	DT ETS	Dampened Trend Exponential Smoothing	Forecast using weighted averages of historical time series that are exponentially decayed and account for seasonality and trends using a "dampening" method to correct for over-forecasting.
Traditional	Linear Trend	Linear Trend Model	Forecast using a trend calculated from historical time series data using linear regression.
Machine Learning	KKNN	K-Nearest Neighbor	Forecast using the "k" most similar case from historical time series data and aggregates them.
Machine Learning	NNAR	Neural Network Autoregression	Forecast using a trend calculated from historical time series data using linear regression.
Machine Learning	XGBOOST	Extreme Gradient Boosting	Forecast using an ensemble of gradient-boosted, regression tree algorithms.
Benchmark	Drift	Drift Method Benchmark	Forecast using the average change of the historical time series data.
Benchmark	Naïve	Naïve Method Benchmark	Forecast using the value of the most recent time series observation.
Benchmark	SNaïve	Seasonal Naïve Method Benchmark	Forecast using the value of the most recent time series observation from the same seasonal period.
Benchmark	Mean	Average Method Benchmark	Forecast using the average value of historical time series data.

adjusting for inflation to make purchasing power comparable at different time periods and the data-normalization procedures of natural logarithm and inverse hyperbolic sine (IHS) transformations. Data normalization involves the mathematical transformation of a variable to make the distribution of data points more closely resemble a normal distribution, the benefits of which are important for statistical reasons. Transforming the distribution to resemble a normal distribution should aid in forecasting accuracy and decrease the impact of outliers within the data set.

The second set of pre-processing steps involves cleaning the data using time series-specific steps. Time series data are data points measuring some phenomena (like sales tax revenue collection) at different past points in time and, as such, are used in forecasting methods. However, time series data require special care as trends (long-term movements in the data) and seasonality (repeating fluctuations in data at regular intervals) are often present and can undermine various methodological procedures. To highlight the importance of accounting for these time series-specific data concerns, we include the following time series-specific pre-processing steps: classical multiplicative time series decomposition procedures to adjust seasonality in monthly and quarterly data in addition to detrending time series data.

Next, we provide a summary of the existing literature on revenue forecasting. Then, we go over the data sources for this manuscript and the pre-processing steps we are undertaking. Next, we present the findings of our forecasting exercise. Finally, we discuss conclusions about the best technique for sales tax forecasting, limitations, and future research areas.

Background

Revenue forecasting in public finance has traditionally fallen into one of two camps: time-series analysis or causal-like approaches (Williams & Calabrese, 2016). Henceforth, these approaches will be referenced as traditional revenue forecasting approaches. The underlying assumption within revenue forecasting is that either previous trends impact future revenues or that a deterministic model of a set of variables can predict revenues. In the case of time series approaches, researchers and practitioners exploit the autocorrelative nature of revenue and the assumption that earlier observed revenue value will help predict future revenues (Williams & Calabrese, 2016). The vast majority of municipalities use a time series approach (Reddick, 2004).

In contrast, causal methods depend on a series of independent variables to predict the dependent forecasted revenue (Mikesell, 2011). Two serious shortcomings of the causal approaches are omitted variable bias (i.e., leaving out important control variables) and a lack of idealized or standardized independent variables (i.e., data that is accessible to local governments). These shortcomings may be why only a small minority of municipalities use this approach (Reddick, 2004).

Machine learning techniques have not been discussed extensively in the public finance literature. According to Chung et al. (2022), 16 articles and conference papers apply machine learning techniques to government revenue forecasts. Of those 16, five were focused on the state level in the United States, and they focused on three states: Indiana, Utah, and Virginia (Carmody & Wiipongwii, 2018; Hansen & Nelson, 1997, 2002; Muh & Jang, 2019; Voorhees, 2006). One omission from the list provided by Chung et al. (2022) was Buxton et al. (2019), which focused on forecasting sales tax revenue in Illinois, broken into ten different retail categories. Findings of the superiority of machine learning or traditional models were not uniform throughout the various studies.

Understanding the accuracy of machine learning techniques as the means of revenue forecasting has not been widely tested nor explored using a wide variety of different study locations. In addition, none of those articles spoke of machine learning techniques being applied actively in practice. Instead, they compared one or more machine learning techniques to traditional models, actual data, existing practices, or other machine learning techniques. A review of the six articles based at the state level found that there has been minimal discussion in those articles of the pre-processing techniques outside of the identification of what pre-processing techniques, if any, were undertaken. In addition, there was no discussion of how variations in pre-processing techniques might change their findings. There is still a wide range of existing knowledge that could be gathered by understanding how both forecasting accuracy and various pre-processing techniques could improve municipal and state revenue forecasting.

The limited literature on revenue forecasting in public finance is not uniform in its account of the superiority of either traditional or machine-learning revenue forecasting techniques. The literature lacks examples of machine learning being applied in practice.

Variation in pre-processing steps also suggests a need for more agreement regarding the best approach.

Data Pre-processing

The literature is mixed regarding the need for pre-processing within revenue forecasting, specifically in cases of machine learning methods. Research has shown that neural networks, when used for forecasting, can adapt to any type of data, thereby negating the need for pre-processing (Gorr, 1994). Similar arguments are made about multivariate autoregressive conditional heteroskedasticity models (Nelson,1996).

In contrast, scholars have argued that pre-processing is necessary, even in the case of machine learning methods, to produce optimal forecasts. Systematic methodology has been proposed to determine whether weights need to be removed during pre-processing (Cottrell et al., 1995). In other cases, neural networks cannot fully capture seasonal or trend variations to reduce forecasting errors and increase efficiency (Zhang & Qi, 2005). Therefore, when using machine learning techniques that depend on neural networks, researchers should either detrend or engage in deseasonalization of their data.

In addition to the question of the need for pre-processing, the research is also mixed on the number and order of necessary pre-processing steps, specifically as it relates to forecasting with time series data (Balkin & Ord, 2000; Miller & Williams, 2004; Zhang & Qi, 2005; Zhang, Cao, & Schniederjans, 2004). This variation is in part due to the fact that not all time series are the same; some have an aspect of seasonality, others may possess an exponential or linear trend, and others may fluctuate around some level or baseline value. Differences in forecasting accuracy in testing three different pre-processing approaches: no special pre-processing, time series differencing, and taking moving averages have been found (Ahmed et al., 2010). In addition, prior research has suggested performing multiple transformations on data, such as log transformation, deseasonalization, and scaling (Ahmed et al., 2010).

One common pre-processing step in revenue forecasting and with time series data that involves currency is adjusting for inflation. Removal of inflation is seen as a key part of the decomposition of revenue and expenditure data (Ammons, 1991, 2001). Armstrong (2001) directly spoke to inflation adjustments within the comprehensive set of principles for forecasting. Further, it has been suggested that “local governments may further benefit by obtaining inflation forecasts from a reputable national firm” (Williams & Kavanagh, 2016, p. 493) and warned that “forecasting the tax revenue or the nominal tax base without adjusting for these factors could lead to significant error” (Williams, 2017, p. 357).

Therefore, due to the need for more evident consensus in pre-processing approaches, researchers would be well served by potentially engaging in a variety of pre-processing steps and varying the order in which they apply them in the case of time series data. As not all time series are the same (Ahmed et al., 2010), what might be the ideal set of pre-processing steps for one revenue forecasting data set might be different for another data set.

Accuracy

Prior research has suggested that picking the correct measurement of error within forecasting can be challenging (Mathews & Diamantopoulos, 1994). In particular, no one measure provides an unambiguous measurement of forecasting performance. In addition, if one relies on multiple measures of accuracy, comparisons among forecasting approaches become difficult (Mathews & Diamantopoulos, 1994). A variety of accuracy measures can be used to determine the accuracy of a forecast, including mean absolute percent error (MAPE), symmetric mean absolute percentage error (sMAPE), revised mean absolute percentage error (r-MAPE), mean squared error (MSE), and model fit.

MAPE provides a percentage of absolute or positive percentage error between the value provided through the forecasting technique and the actual observed value or revenue dollars collected. At the same time, sMAPE adjusts the percentage error calculation to include the sum of the absolute actual observed value and the absolute forecasted value divided by two, which creates a lower and upper bound. r-MAPE divides the traditional MAPE calculation by the number of periods considered. MSE uses the average of the square of the difference between the actual and forecasted value. The model fit looks at statistical measures of the overall fit of the analysis, such as adjusted R-square in the case of linear regression. Each of these approaches has strengths and weaknesses, and it is not appropriate on all occasions. For example, with some of the traditional regression approaches, model fit is a straightforward measure of accuracy. In contrast, MSE depends on the number of observations (Chung et al., 2022).

One argument in the literature is that the sMAPE is the superior accuracy measure (Chung et al., 2022). One of the main reasons for this argument is that in times of variation of the scale of observations between series, a few series with larger values can dominate the comparison within the MAPE (Chatfield, 1988). Equal errors above the actual values and equal errors below the actual value will not create the same absolute percent error (APE) (Makridakis, 1993). Equal values above lead to a greater APE (Armstrong & Collopy, 1992). Therefore, there are numerous examples of the sMAPE being used or recommended (Chung et al., 2022; Hyndman & Koehler, 2006; Makridakis & Hibon, 2000; Taieb et al., 2012; Williams & Miller, 1999).

In contrast, a separate line of argument in the literature points to errors with sMAPE (Goodwin & Lawton, 1999). For example, in some instances, a non-monotonic relationship can occur between sMAPE and absolute forecasting errors. Inconsistent performance in sMAPE estimators has been found depending on whether the forecasted value underestimated or overestimated the actual value (Tayman & Swanson, 1999). In addition, they found that differences between sMAPE and MAPE were related to the side of the over or underestimate. In addition, sMAPE tends towards high error values when the error is small (Mathai et al., 2016). Similar to sMAPE, there are numerous examples of MAPE being used or recommended within the literature (Callen et al., 1996; Halimawan & Sukarno, 2013; Prayudani et al., 2019; Singh, Hussain, & Bazaz, 2017; Vivas, Allende-Cid, & Salas, 2020).

Therefore, the literature has not coalesced around one particular measurement tool for forecasting accuracy. This lack of uniformity harkens back to Mathews and Diamantopoulos's (1994) argument that no one measure provides an unambiguous measurement of performance. Therefore, a measurement of accuracy must be selected and justified.

Texas Sales Tax

Similar to prior studies, the basis of our study will be at the state level, in part due to variations in sales tax laws between states in the United States. However, we are not following the prior precedent of using Illinois, Indiana, Utah, and Virginia. Texas was selected because of the large number of cities that issue sales taxes and the relatively limited variation in sales tax rates. For our sample, we collected data on over 1,000 cities for 16 years. This gives us a substantial number of jurisdictions to assess forecast methods. Additionally, sales tax rates are relatively stable in Texas, both between and within cities.

In Texas, counties and cities can impose a 1% sales tax, with the option to add another 1% through various entities like economic development corporations, public transportation governments, and police and fire districts. The state has a maximum sales tax rate of 8.25%, with 6.25% allocated to the state's general fund (Texas Comptroller of Public Accounts, 2015). Registered businesses are required to collect sales taxes on behalf of the state comptroller's office during normal operations and to submit sales tax revenue to the state on a monthly, quarterly, or annual basis, depending on the size of their organizations. Using only Texas, we ensure that institutional rules on collection, tax base, and other state administrative policies are uniform across the cross-sectional units.

Forecasting Methodology and Comparisons

To understand the most accurate forecasting approaches, monthly sales tax collection data for every city in Texas were collected between January 1991 and December 2017 using the Texas Comptroller website. Unfortunately, not all the monthly data were available for the entire time frame for every city in Texas. Only cities with complete time series were included in the analyses, resulting in 822 cities with complete monthly data, 976 cities with complete quarterly data, and 1,005 cities with complete yearly data. While rare, some cities in Texas changed their sales tax rates during the period under study. Cities that did not have a uniform sales tax throughout the period of study were time-series comparable. This ensured that the rate changes would not affect the forecasting accuracy.

To compare forecasts, we ran each pre-processing step and forecasting model on each city using its monthly, quarterly, and yearly data. This approach resulted in a unique forecast for each city. Forecast accuracy was evaluated using the MAPE score on the latest 24 months, eight quarters, or four years of data. These city-level MAPE scores were then aggregated for every model-preprocessing (MP) combination, producing MAPE averages and standard deviations at the MP level. City-level forecasts three times the interquartile range were identified as outliers and removed before MP aggregation.

Results

To evaluate and compare MP forecasting performance, we have provided the average MAPE, standard deviation, performance ranking, the percentage difference from the best forming model MAPE called diminished performance (%), the minimum and maximum MAPE within each model, the total number of cities forecasted in each model, and the number of outlier cities removed. These metrics provide useful context on how models and pre-processing steps perform across the entire population of forecasted cities. To streamline the results, we only included the top 30 MP forecasts for the monthly and quarterly forecasts and every forecast that performed above the MAPE average for the yearly forecasts. Monthly MP forecasts are in Table 2,

Table 2. Monthly Sales Tax Forecast Accuracy

	Average	Standard Deviation	Min	Max	Performance Ranking	Diminished Performance	Forecasts	Outlier Forecasts
(IHS) (Detrend)								
KKNN	1.000	0.452	0.276	2.346	1	--	759	63
XGBOOST	1.044	0.508	0.336	2.443	2	4.4%	756	66
SNaïve*	1.128	1.037	0.117	8.293	5	12.8%	822	0
Linear Trend	1.192	0.431	0.472	2.318	6	19.1%	747	75
Mean*	1.381	0.777	0.476	6.782	15	38.1%	822	0
Naïve*	1.644	1.502	0.462	22.468	21	64.4%	822	0
Drift*	1.660	1.538	0.469	23.209	22	66.0%	822	0
NNAR	2.763	1.902	0.286	8.890	39	176.2%	756	66
(Log) (Detrend)								
KKNN	1.075	0.501	0.289	2.544	3	7.5%	764	58
XGBOOST	1.081	0.518	0.275	2.571	4	8.1%	745	77
SNaïve*	1.208	1.125	0.123	9.041	7	20.8%	822	0
Linear Trend	1.273	0.473	0.501	2.509	11	27.3%	750	72
Mean*	1.474	0.846	0.505	7.456	17	47.4%	822	0
Naïve*	1.756	1.621	0.479	23.987	27	75.6%	822	0
Drift*	1.773	1.659	0.487	24.779	28	77.2%	822	0
NNAR	2.885	1.977	0.249	9.119	40	188.4%	749	73
(IHS)								
DT ETS	1.213	0.752	0.203	3.456	8	21.3%	759	63
ARIMA	1.226	0.678	0.257	3.286	9	22.6%	747	75
SNaïve*	1.610	1.688	0.170	30.373	20	61.0%	822	0
Linear Trend	1.821	0.896	0.434	4.321	31	82.1%	750	72
Drift*	1.834	1.781	0.428	24.261	32	83.3%	822	0
Naïve*	1.840	1.731	0.491	23.288	33	83.9%	822	0
(IHS) (SA)								
KKNN	1.255	0.716	0.181	3.361	10	25.4%	760	62
XGBOOST	1.466	0.665	0.302	3.374	16	46.5%	753	69
SNaïve*	1.608	1.687	0.170	30.430	19	60.8%	822	0
Naïve*	1.676	1.812	0.205	22.710	23	67.6%	822	0
Linear Trend	1.753	1.093	0.185	4.898	26	75.3%	776	46
Drift*	1.812	1.879	0.195	23.956	30	81.2%	822	0
NNAR	2.986	2.130	0.288	10.274	41	198.5%	747	75
(Log)								
DT ETS	1.307	0.820	0.216	3.912	12	30.6%	759	63
ARIMA	1.313	0.738	0.267	3.557	13	31.2%	749	73
SNaïve*	1.722	1.841	0.177	33.702	25	72.2%	822	0
Linear Trend	1.954	0.979	0.451	4.707	36	95.3%	754	68
Drift*	1.959	1.923	0.445	25.883	37	95.9%	822	0
Naïve*	1.965	1.870	0.510	24.846	38	96.5%	822	0
(Log) (SA)								
KKNN	1.337	0.771	0.189	3.607	14	33.7%	760	62
XGBOOST	1.548	0.717	0.304	3.597	18	54.7%	754	68
SNaïve*	1.721	1.841	0.177	33.773	24	72.0%	822	0
Naïve*	1.793	1.956	0.215	24.566	29	79.3%	822	0
Linear Trend	1.873	1.177	0.192	5.296	34	87.2%	777	45
Drift*	1.938	2.027	0.205	25.913	35	93.8%	822	0

*Denotes a benchmark model.
Grey highlighted row = ranked 1-5
Green highlighted row = ranked 6-10

Table 3. Quarterly Sales Tax Forecast Accuracy

	Average	Standard Deviation	Min	Max	Performance Ranking	Diminished Performance	Forecasts	Outlier Forecasts
(IHS) (Detrend)								
Linear Trend	0.406	0.261	0.040	1.226	1	--	893	83
KKNN	0.474	0.312	0.027	1.472	3	16.5%	900	76
Mean*	0.544	0.594	0.040	7.161	5	33.9%	976	0
XGBOOST	0.607	0.206	0.223	1.190	7	49.3%	840	136
SNaïve*	0.706	0.835	0.029	12.133	9	73.8%	976	0
Naïve*	0.833	1.027	0.056	12.735	11	105.0%	976	0
Drift*	0.852	1.055	0.058	13.141	12	109.5%	976	0
NNAR	2.506	2.577	0.077	11.811	41	516.4%	893	83
(Log) (Detrend)								
Linear Trend	0.431	0.279	0.042	1.314	2	6.1%	893	83
KKNN	0.502	0.333	0.028	1.571	4	23.5%	899	77
Mean*	0.580	0.637	0.041	7.605	6	42.6%	976	0
XGBOOST	0.634	0.232	0.208	1.319	8	55.9%	847	129
SNaïve*	0.753	0.897	0.030	12.953	10	85.2%	976	0
Naïve*	0.887	1.101	0.059	13.689	14	118.3%	976	0
Drift*	0.907	1.131	0.062	14.126	15	123.1%	976	0
NNAR	2.518	2.595	0.071	11.947	42	519.4%	879	97
(IHS)								
DT ETS	0.875	0.554	0.086	2.574	13	115.2%	367	36
DT ETS	0.923	0.649	0.080	2.897	16	127.0%	540	33
ARIMA	0.938	0.641	0.097	2.973	17	130.7%	902	74
Naïve*	1.196	1.428	0.060	16.888	24	194.3%	976	0
Drift*	1.225	1.466	0.096	16.863	26	201.4%	976	0
SNaïve*	1.333	1.518	0.058	25.048	32	227.8%	976	0
Linear Trend	1.536	1.125	0.077	4.981	38	277.9%	930	46
(Log)								
DT ETS	0.943	0.609	0.090	2.846	18	131.9%	367	30
DT ETS	0.975	0.691	0.085	3.015	19	140.0%	542	37
ARIMA	0.998	0.690	0.101	3.184	21	145.5%	903	73
Naïve*	1.274	1.534	0.062	18.449	28	213.3%	976	0
Drift*	1.304	1.574	0.100	18.081	30	220.9%	976	0
SNaïve*	1.419	1.633	0.061	27.341	35	249.0%	976	0
Linear Trend	1.622	1.188	0.081	5.280	39	299.1%	928	48
(IHS) (SA)								
KKNN	0.979	0.683	0.088	3.087	20	140.7%	906	70
Naïve*	1.192	1.428	0.058	16.941	23	193.1%	976	0
Drift*	1.221	1.458	0.050	16.536	25	200.4%	976	0
SNaïve*	1.332	1.518	0.058	25.068	31	227.8%	976	0
XGBOOST	1.411	0.668	0.174	3.136	33	247.1%	904	72
Linear Trend	1.534	1.134	0.066	5.019	37	277.4%	931	45
(Log) (SA)								
KKNN	1.044	0.737	0.092	3.352	22	156.9%	908	68
Naïve*	1.269	1.534	0.061	18.511	27	212.2%	976	0
Drift*	1.300	1.566	0.053	17.731	29	219.9%	976	0
SNaïve*	1.418	1.634	0.061	27.366	34	248.9%	976	0
XGBOOST	1.464	0.716	0.171	3.332	36	260.2%	907	69
Linear Trend	1.624	1.202	0.069	5.286	40	299.5%	930	46

*Denotes a benchmark model.
Grey highlighted row = ranked 1-5
Green highlighted row = ranked 6-10

Table 4. Annual Sales Tax Forecast Accuracy

	Average	Standard Deviation	Min	Max	Performance Ranking	Diminished Performance	Forecasts	Outlier Forecasts
(IHS)								
Drift*	1.400	1.771	0.032	20.717	1	--	1005	0
ARIMA	1.437	0.951	0.060	4.336	2	2.6%	942	63
Naïve*	1.447	1.609	0.048	17.739	3	3.4%	1005	0
DT ETS	1.489	1.335	0.040	5.886	5	6.4%	939	63
Linear Trend	1.511	1.145	0.045	5.286	6	8.0%	951	54
NNAR	1.995	1.857	0.028	8.756	11	42.5%	902	103
Mean*	3.627	2.920	0.053	19.969	13	159.1%	1005	0
(Log)								
Drift*	1.480	1.888	0.033	22.278	4	5.8%	1005	0
ARIMA	1.512	1.003	0.063	4.574	7	8.0%	941	64
Naïve*	1.529	1.714	0.051	19.076	8	9.2%	1005	0
DT ETS	1.570	1.413	0.041	6.208	9	12.2%	941	63
Linear Trend	1.581	1.192	0.047	5.559	10	13.0%	948	57
NNAR	2.130	2.000	0.038	9.143	12	52.2%	904	101
Mean*	3.822	3.083	0.055	21.167	14	173.1%	1005	0

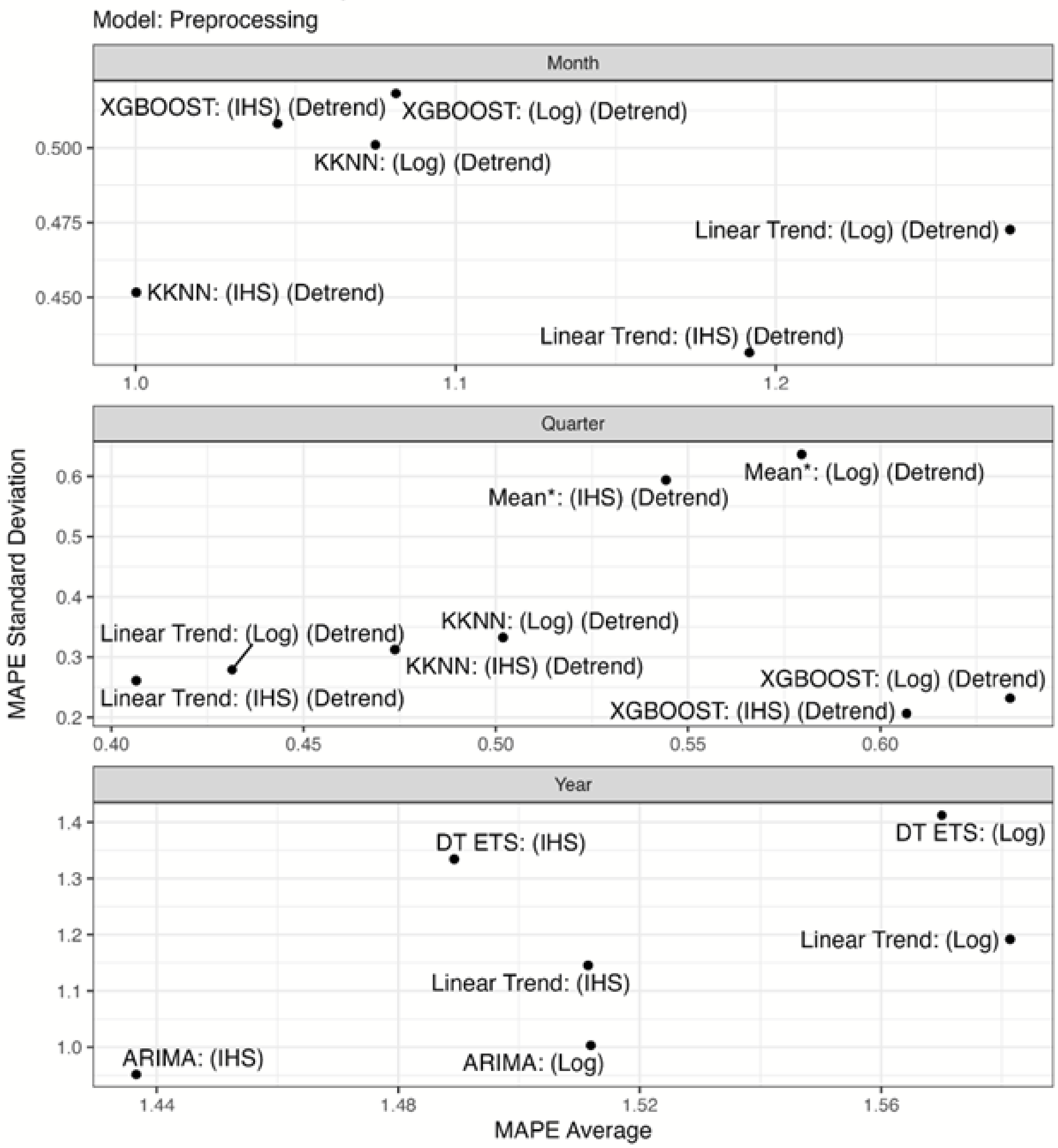
*Denotes a benchmark model
Grey highlighted row = ranked 1-5

quarterly forecasts are in Table 3, and annual forecasts are in Table 4. The complete results are in the appendix, which is available on the journal's website.

To identify the most accurate MP forecasts using both the average MAPE and standard deviation MAPE, k-means exploratory cluster analysis was performed. K-means exploratory cluster analysis groups similar observations together into groups based on multiple variables, which will give us an approximation of both the most accurate (measured using average MAPE) and the most consistent (measured using standard deviation of MAPE) models and pre-processing steps. As the goal of forecasting may be accuracy or consistency, this provides interesting findings about the differences between the two based on the approach. The optimal number of clusters was determined using a visual inspection of the explanatory power of different numbers of clusters called the elbow method (Humaira & Rasyidah, 2020), and it indicated that eight clusters from the monthly forecasts, four clusters from the quarterly, and four clusters from the yearly forecasts were optimal. The most accurate MP forecasts from the cluster analyses are in Figure 1.

The performance of each MP average MAPE is compared to the most accurate MP forecast and reported in the column Diminished Performance (%). For the monthly forecasts, the most accurate model had an average MAPE of 1 using KKNN with IHS detrended data. The average diminished performance of the top ten MP forecasts was 15.78%, and for all models presented in Table 1, it was 44.97%. For the quarterly forecasts, the most accurate model was Linear Trend with IHS detrended data that had an average MAPE of 0.406. The average diminished performance of the top 10 quarterly MP forecasts was 42.98%, and for all models presented in Table 2, it was 123.66%. For the yearly forecasts, the most accurate model was the Drift benchmark model with IHS data with an average MAPE of 1.400. The average diminished performance of the top five MP forecasts was 4.54%, and for all models presented in Table 3, it

Figure 1. Most Accurate Yearly Forecasts



was 38.1%. Interestingly, the most accurate forecasts from each period used IHS pre-processing steps, while a different model resulted in more accurate forecasts for each period. This suggests that stable forecasting improvements can be made across different time periods with the implementation of similar pre-processing steps, but not through similar modeling approaches.

Table 5 presents the results of averaged percent diminished MP performance across different models but holds pre-processing steps constant. Table 6 compares the average

Table 5. Average Percentage Diminished Performance with Constant Preprocessing

	All Forecasts				No Outlying Forecasts	
	Average Diminished Performance	**Average Performance Ranking**	**Total Models**	**Outlier Models**	**Average Diminished Performance**	**Average Performance Ranking**
Monthly						
(IHS) Detrend)	47.6%	1	8	0	47.6%	1
(Log) Detrend)	56.5%	2	8	0	56.5%	2
(IHS)	96.1%	3	7	0	96.1%	3
(IHS) (SA)	108.4%	4	8	0	108.4%	4
(Log)	109.6%	5	7	0	109.6%	5
(Log) (SA)	123.3%	6	8	0	123.3%	6
(Detrend)	1,672.0%	7	8	5	1,094.5%	7
No Transformation	1,955.2%	8	7	4	1,454.1%	9
(SA)	2,097.0%	9	8	5	1,406.9%	8
Quarterly						
(IHS) Detrend)	113.1%	1	8	0	113.1%	1
(Log) Detrend)	121.8%	2	8	0	121.8%	2
(IHS)	266.9%	3	8	0	266.9%	3
(Log)	289.9%	4	8	0	289.9%	4
(IHS) (SA)	353.8%	5	8	0	353.8%	5
(Log) (SA)	386.1%	6	8	0	386.1%	6
(Detrend)	2,452.2%	7	8	1	1,798.5%	7
No Transformation	3,862.2%	8	8	4	2,894.3%	8
(SA)	4,721.9%	9	8	5	3,062.8%	9
Yearly						
(IHS)	31.7%	1	7	0	31.7%	1
(Log)	39.1%	2	7	0	39.1%	2
No Transformation	1,440.5%	3	7	0	1,440.5%	3

diminishing performance of different pre-processing steps but holds the model constant. Table 7 provides the range of improvement from pre-processing (Table 5), and modeling (Table 6).

Discussions

The purpose of this study is to identify the most accurate sales tax forecasting method, and the analysis suggests three overarching lessons: pre-processing makes the most significant difference in forecasting accuracy, understanding the unique characteristics of time series data improves forecasting performance, and modeling choices matter, but less than the prior literature and practice suggested.

Processing Matters

The most significant improvement in our analysis occurred when data were transformed using the IHS method, followed closely by logging–both of which were data-normalizing pre-processing steps. This finding is consistent across all three periods. Most studies recommend some form of pre-processing that normalizes the time series data, including previous studies on sales tax forecasting (Williams & Calabrese, 2016). What this study suggests is that regardless of

Table 6. Average Percentage Diminished Performance with Constant Models

	All Forecasts				No Outlying Forecasts	
	Average Diminished Performance	**Average Performance Ranking**	**Total Preprocessing**	**Outlier Models**	**Average Diminished Performance**	**Average Performance Ranking**
Monthly						
XGBOOST	395.8%	1	6	0	395.8%	6
KKNN	396.5%	2	6	0	396.5%	7
ARIMA	439.2%	3	3	0	439.2%	8
DT ETS	461.0%	4	3	0	461.0%	9
Linear Trend	565.0%	5	9	0	565.0%	10
SNaïve*	613.4%	6	9	3	49.9%	1
Naïve*	696.5%	7	9	3	77.9%	2
Drift*	714.6%	8	9	3	82.9%	3
Mean*	1,090.9%	9	9	3	233.3%	5
NNAR	1,231.1%	10	6	2	197.6%	4
Quarterly						
KKNN	729.1%	1	6	0	729.1%	5
XGBOOST	780.6%	2	6	0	780.6%	7
DT ETS	935.6%	3	6	0	935.6%	8
ARIMA	955.1%	4	3	0	955.1%	9
Linear Trend	1,119.9%	5	9	0	1,119.9%	10
Naïve*	1,161.6%	6	9	2	502.2%	2
Drift*	1,184.4%	7	9	2	515.1%	3
SNaïve*	1,250.5%	8	9	2	447.3%	1
Mean*	2,406.8%	9	9	2	749.6%	6
NNAR	3,149.4%	10	6	2	623.1%	4
Yearly						
ARIMA	383.5%	1	3	0	383.5%	1
Linear Trend	398.4%	2	3	0	398.4%	2
Drift*	416.0%	3	3	0	416.0%	3
Naïve*	422.2%	4	3	0	422.2%	4
DT ETS	426.9%	5	3	0	426.9%	5
NNAR	586.8%	6	3	0	586.8%	6
Mean*	892.3%	7	3	0	892.3%	7

*Denotes a benchmark model.

the period used, the biggest improvement in forecast performance comes from IHS transformation.

Two exciting patterns emerged in the comparisons of pre-processing steps. The first pattern was that removing outlying average forecasts did not drastically alter the rank order of model performance. In the monthly and quarterly forecasts, the first seven of nine pre-processing steps stayed in the same rank order regardless of whether outlying model forecasts were included or removed. This suggests that improvements to individual city forecasts from data pre-processing are likely to improve forecasts even in cities that are relative forecasting outliers.

The second pattern was the lower average diminishing performance of the top six pre-processing steps in the monthly and quarterly data, as seen in Table 6. The low average diminished performance associated with data pre-processing is stark compared to the much larger average diminished performance across models (i.e., Table 7). This finding suggests that variations within a pre-processing step and across models are smaller than vice versa. Making sure that data are transformed results in better and more stable forecast accuracy.

Table 7. Range of Improvement from Preprocessing

	All Forecasts		No Outlying Forecasts	
	Model Diminished Performance Range	**Preprocessing Diminished Performance Range**	**Model Diminished Performance Range**	**Preprocessing Diminished Performance Range**
Monthly	835.3%	2,049.3%	515.0%	1,406.5%
Quarterly	2,420.3%	4,608.8%	672.6%	2,949.7%
Yearly	508.8%	1,408.8%	508.8%	1,408.8%

Table 8. Inflation Adjustment Performance Comparison

					% of Forecasts Better/Worse Adjusting for Inflation	
	Average	**Standard Deviation**	**Min**	**Max**	**Better**	**Worse**
Monthly	2.4%	3.2%	-2.7%	10.0%	85.4%	14.6%
Quarterly	1.4%	4.6%	-22.2%	12.5%	92.9%	7.1%
Yearly	31.3%	67.7%	-9.8%	193.6%	57.1%	42.9%

A standard pre-processing step in time series data is adjusting for inflation, which is important for integrating[1] time series data. Table 8 shows the average improvement in accuracy for adjusting a forecast for inflation. The average improvement from inflation adjustment is 2.4%, 1.4%, and 31.3% for monthly, quarterly, and yearly periods, respectively. While there is forecasting improvement in all three periods, the improvement in average forecasting accuracy is considerably smaller than improvements gained by IHS or logging the time series data. Another interesting development is that adjusting for inflation did not always improve forecasting accuracy. Adjusting for inflation led to reductions in model/pre-processing accuracy of 14.6%, 7.1%, and 42.9% for monthly, quarterly, and yearly data, respectively. Adjusting for inflation is an important part of forecasting financial data, but it does not always translate to improved forecasting accuracy.

Understanding Time-Series Data Matters

The results highlight the importance of understanding the nature of time series data and using that knowledge when forecasting. In our analyses, detrending and seasonally adjusting data improved forecasting accuracy. The top seven performing models in the monthly data and all top 10 models in the quarterly data were detrended time series. Adjusting for seasonality did not result in the same performance improvements, but it did result in relative improvements in accuracy over a non-seasonally adjusted time series. Further, accounting for seasonality and trends resulted in improvements in machine learning models, which were the best-performing models in the monthly and quarterly forecasts.

Finally, another crucial finding was that the time interval affects the effectiveness of models. The best-performing models change across time intervals, and certain pre-processing steps make a bigger difference for certain time periods. For example, seasonally adjusting data does not apply to annual data, but adjusting for inflation made the biggest positive difference to

[1] Integration is a statistical term that refers broadly to "detrending" time series data and making it more statistically "stable."

yearly data by a large percentage. What time interval is best? The answer depends on what data is most thorough and accurately reflects when the sales tax was collected (Overton, Nukpezah, & Ismayilov, 2017). There were significant performance differences between the three periods, suggesting that attention should be paid to the frequency of time series data used for sales tax forecasts, with a preference for the intervals at which the sales tax was collected.

Modeling Matters, But Not as Much as Everything Else

Table 7 shows that forecasting accuracy improved most from pre-processing decisions rather than modeling decisions. In addition, Table 7 illustrates how the average diminished performance based on modeling decisions was much larger than that of the pre-processing steps. Pre-processing steps led to considerably larger improvements in forecasting accuracy than modeling decisions. Therefore, the focus of both public finance professionals and researchers should be on conducting the appropriate pre-processing steps.

One additional pattern that emerged is that the rank of average model forecasting accuracy changed drastically once outlying average forecasts were removed. Before outlier removal, XGBOOST, KNN, ARIMA, DT ETS, and Linear Trend models performed best. However, once outliers were removed, the average performance of all four benchmark models outperformed that of the machine learning and traditional methods. However, it should be noted that the XGBOOST, KNN, ARIMA, DT ETS, and Linear Trend models did not produce any model-level MAPE outliers, suggesting that their accuracy is less prone to variations in pre-processing. Therefore, one should carefully weigh average forecasting performance against the likelihood of generating forecasting outliers when evaluating forecast-generating methods.

Conclusion

Returning to the question we started with, what is the best technique for forecasting sales tax revenue, the answer is not as simple, as traditional methods can outperform machine learning techniques or vice versa. In fact, the existing literature tackling this problem might have been asking the wrong question in its entirety. Scholars such as Chung et al. (2022) and Buxton et al. (2019) provided direct comparisons of machine learning techniques with traditional ones regarding forecasting accuracy. They should consider additional analysis to determine whether their findings hold in the case of different pre-processing steps. In our attempt to answer the question, findings held that perhaps the question should have been: What pre-processing steps did you take before engaging in forecasting? In many ways, we provide strong evidence that the arguments by Gorr (1994) and Nelson (1996) are fundamentally flawed while supporting the findings of Zhang and Qi (2005) that neural networks cannot fully capture seasonal or trend variation and pre-processing of time series data matters.

However, not all pre-processing steps are as important as the previous literature suggested (see Ammons, 1991, 2001; Armstrong, 2001; Williams & Kavanagh, 2016). Adjusting for inflation is an important part of forecasting, but it does not always translate into improved forecasting accuracy. Therefore, researchers and practitioners should exercise caution in adjusting for inflation. In contrast, IHS, or logging the time series data, led to larger gains in overall forecast accuracy.

Understanding the nature of time series data is essential for individuals engaging in revenue forecasting. Detrending and seasonally adjusting data should be common practice for those working with monthly or quarterly data. In addition, the interval affects the effectiveness of models, and individuals should attempt to select what accurately reflects when the sales tax was collected.

So, what is the best technique for forecasting sales tax revenue? Should local governmental officials study machine learning techniques to improve their forecasting accuracy? Well, part of the answer to the question relies on outliers. XGBOOST, KNN, ARIMA, DT ETS, and Linear Trend models did not produce any model-level MAPE outliers, suggesting that their accuracy is less prone to variations in pre-processing. However, before outlier removal, the XGBOOST, KNN, ARIMA, DT ETS, and Linear Trend models performed best. Which individual approach performed best–machine learning or traditional models–depended on the interval of time. Therefore, a rush to machine learning techniques may not be necessary for municipalities if they collect sales taxes quarterly or yearly.

Instead of focusing on the technique, revenue forecasting practitioners should be focusing on the pre-processing steps they are using on their data. The pre-processing provided much better forecasting accuracy than simply a model selection. Therefore, future researchers should shift the focus from machine learning or traditional models towards an approach that takes a holistic approach and includes various pre-processing steps. Testing the findings from this research on data from a different state with different rules on collection, tax base, and other state administrative policies would be a logical next step toward understanding and expanding the knowledge base of revenue forecasting. Regardless, the focus within the literature and in practice should shift from what forecasting technique performs best to what pre-processing steps in combination with a forecasting technique perform best. Simply focusing on technique ignores the larger concern in achieving greater accuracy–pre-processing steps.

Disclosure Statement

The authors declare that there are no conflicts of interest that relate to the research, authorship, or publication of this article.

References

Ahmed, N. K., Atiya, A. F., Gayar, N. E., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, *29*(5-6), 594-621. https://doi.org/10.1080/07474938.2010.481556

Ammons, D. N. (1991). *Administrative analysis for local government: Practical application of selected techniques*. Carl Vinson Institute of Government.

Ammons, D. N. (2001). *Tools for decision making: A practical guide for local government*. CQ Press.

Armstrong, J. S. (2001). Selecting forecasting methods. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners* (pp. 365–386). Kluwer Academic.

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, *8*(1), 69-80. https://doi.org/10.1016/0169-2070(92)90008-W

Balkin, S. D., & Ord, J. K. (2000). Automatic neural network modeling for univariate time series. *International Journal of Forecasting*, *16*(4), 509-515. https://doi.org/10.1016/S0169-2070(00)00072-8

Boyd, D. J., & Dadayan, L. (2014). *State tax revenue forecasting accuracy*. Rockefeller Institute of Government.

Bretschneider, S., & Gorr, W. (1992). Economic, organizational, and political influences on biases in forecasting state sales tax receipts. *International Journal of Forecasting*, *7*(4), 457-466. https://doi.org/10.1016/0169-2070(92)90029-9

Bretschneider, S. I., Gorr, W. L., Grizzle, G., & Klay, E. (1989). Political and organizational influences on the accuracy of forecasting state government revenues. *International Journal of Forecasting*, *5*(3), 307-319. https://doi.org/10.1016/0169-2070(89)90035-6

Buettner, T., & Kauder, B. (2010). Revenue forecasting practices: Differences across countries and consequences for forecasting performance. *Fiscal Studies*, *31*(3), 313-340. https://doi.org/10.1111/j.1475-5890-2010.000117.x

Buxton, E., Kriz, K., Cremeens, M., & Jay, K. (2019). An auto-regressive deep learning model for sales tax forecasting from multiple short time series. In M. A. Wanif, T. M. Khoshgoftaar, D. Wang, H. Wang, & N. Seliya (Eds.), *Proceedings of the 18th IEEE international conference on machine learning and applications* (pp. 1359-1364). IEEE. https://doi.org/10.1109/ICMLA.2019.00221

Callen, J. L., Kwan, C. C., Yip, P. C., & Yuan, Y. (1996). Neural network forecasting of quarterly accounting earnings. *International Journal of Forecasting*, *12*(4), 475-482. https://doi.org/10.1016/S0169-2070(96)00706-6

Carmody, K., & Wiipongwii, T. (2018). *Virginia tax revenue forecasting* [Unpublished Manuscript]. Department of Public Policy, College of William and Mary.

Chatfield, C. (1988). Apples, oranges and mean square error. *International Journal of Forecasting*, *4*(4), 515-518. https://doi.org/10.1016/0169-2070(88)90127-6

Chung, I. H., Williams, D. W., & Do, M. R. (2022). For better or worse? Revenue forecasting with machine learning approaches. *Public Performance & Management Review*, *45*(5), 1133-1154. https://doi.org/10.1080/15309576.2022.2073551

Cole, R. (1969). Data errors and forecasting accuracy. In J. A. Mincer (Ed.), *Economic forecasts and expectations: analysis of forecasting behavior and performance* (pp. 47-82). National Bureau of Economic Research.

Cottrell, M., Girard, B., Girard, Y., Mangeas, M., & Muller, C. (1995). Neural modeling for time series: A statistical stepwise method for weight elimination. *IEEE Transactions on Neural Networks*, *6*(6), 1355-1364. https://doi.org/10.1109/72.471372

Frank, H. A. (1990). Municipal revenue forecasting with time. Series models: A Florida case study. *American Review of Public Administration*, *20*(1), 45-59. https://doi.org/10.1177/027507409002000104

Goodwin, P., & Lawton, R. (1999). On the asymmetry of the symmetric MAPE. *International Journal of Forecasting*, *15*(4), 405-408. https://doi.org/10.1016/S0169-2070(99)00007-2

Gorr, W. L. (1994). Research prospective on neural network forecasting. *International Journal of Forecasting*, *10*(1), 1-4. https://doi.org/10.1016/0169-2070(94)90044-2

Grizzle, G. A., & Klay, W. E. (1994). Forecasting state sales tax revenues: Comparing

the accuracy of different methods. *State & Local Government Review*, *26*(3), 142-152.
Halimawan, A. A., & Sukarno, S. (2013). Stock price forecasting accuracy analysis using mean absolute deviation (MAD) and mean absolute percentage error (MAPE) on smoothing moving average and exponential moving average indicator (Empirical study 10 LQ 45 stock with largest capitalization from period Feb-Jul 2013). *Indonesian Journal of Business Administration, 2*(13), 68283.
Hansen, J. V., & Nelson, R. D. (1997). Neural networks and traditional time series methods: A synergistic combination in state economic forecasts. *IEEE Transactions on Neural Networks*, *8*(4), 863-873. https://doi.org/10.1109/72.595884
Hansen, J. V., & Nelson, R. D. (2002). Data mining of time series using stacked generalizers. *Neurocomputing, 43*(1-4), 173-184. https://doi.org/10.1016/S0925-2313(00)00364-7
Humaira, H., & Rasyidah, R. (2020). Determining the appropriate cluster number using elbow method for *k*-means algorithm. In R. Hidayat, Y. Sonatha, A. S. Ahmar, J. Simarmata, & D. Abdullah (Eds.), *Proceedings of the 2nd workshop on multidisciplinary and applications (WMA) 2018*. EAI. https://doi.org/10.4108/eai.24-1-2018.2292388
Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679-688. https://doi.org/10.1016/j.ilforecast.2006.03.001
Makridakis, S. (1993). Accuracy measures: Theoretical and practical concerns. *International Journal of Forecasting*, *9*(4), 527-529. https://doi.org/10.1016/0169-2070(93)90079-3
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
Mathai, A. V., Agarwal, A., Angampalli, V., Narayanan, S., & Dhakshayani, E. (2016). Development of new methods for measuring forecast error. *International Journal of Logistics Systems and Management*, *24*(2), 213-225. https://doi.org/10.1504/IJLSM.2016.076472
Mathews, B. P., & Diamantopoulos, A. (1994). Towards a taxonomy of forecast error measures: A factor-comparative investigation of forecast error dimensions. *Journal of Forecasting*, *13*(4), 409-416. https://doi.org/10.1002/for.3980130406
Mikesell, J. (2011). *Fiscal administration: Analysis and applications for the public sector* (8th edition). Wadsworth.
Miller, D. M., & Williams, D. (2004). Damping seasonal factors: Shrinkage estimators for the X–12–ARIMA program. *International Journal of Forecasting*, *20*(4), 529-549. https://doi.org/10.1016/j.ijforecast.2004.03.002
Muh, K. L., & Jang, S. B. (2019). A time-series data prediction using tensorflow neural network libraries. *KIPS Transactions on Computer and Communication Systems*, *8*(4), 79-86. https://doi.org/10.3745/KTCCS.2019.8.4.79
Nelson, D. B. (1996). Asymptotic filtering theory for multivariate ARCH models. *Journal of Econometrics*, *71*(1-2), 1-47. https://doi.org/10.1016/0304-4076(94)01679-8
Overton, M., Nukpezah, J. A., & Ismayilov, O. (2017). Prepayments, late payments, and sales tax revenue volatility in Texas cities. *Public Money & Management*, *37*(7), 469-476. https://doi.org/10.1080/09540962.2017.1372099
Prayudani, S., Hizriadi, A., Lase, Y. Y., & Fatmi, Y. (2019). Analysis accuracy of forecasting measurement technique on random K-nearest neighbor (RKNN) using MAPE and MSE.

Journal of Physics: Conference Series, *1361*(1), 012809. https://doi.org/10.1088/1742-6596/1361/1/012089
Reddick, C. G. (2004). An empirical examination of revenue forecasting techniques in local governments. *Municipal Finance Journal*, *24*(4), 25-48.
Rodgers, R., & Joyce, P. (1996). The effect of underforecasting on the accuracy of revenue forecasts by state governments. *Public Administration Review*, *56*(1), 48-56. https://doi.org/10.2307/3110053
Rubin, M., Mantell, N., & Pagano, M. A. (1999). Approaches to revenue forecasting by state and local governments. *Proceedings of the Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, *92*, 205-221.
Singh, S., Hussain, S., & Bazaz, M. A. (2017). Short term load forecasting using artificial neural network. In P. K. Gupta, A. K. Singh, V. Tyagi, & S. P. Ghrera (Eds.), *2017 Fourth international conference on image information processing (ICIIP)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICIIP.2017.8313703
Taieb, S. B., Bontempi, G., Atiya, A. F., & Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Systems with Applications*, *39*(8), 7067-7083. https://doi.org/10.1016/j.eswa.2012.01.039
Tayman, J., & Swanson, D. A. (1999). On the validity of MAPE as a measure of population forecast accuracy. *Population Research and Policy Review*, *18*, 299-322. https://doi.org/10.1023/A;100616641805
Texas Comptroller of Public Accounts. (2015). *Sources of revenue: A history of state taxes and fees in Texas.* Texas Comptroller of Public Accounts.
Vivas, E., Allende-Cid, H., & Salas, R. (2020). A systematic review of statistical and machine learning methods for electrical power forecasting with reported MAPE score. *Entropy*, *22*(12), 1412. https://doi.org/10.10.3390/e22121412
Voorhees, W. R. (2006). Neural networks and revenue forecasting: A smarter forecast? *International Journal of Public Policy*, *1*(4), 379-388. https://doi.org/10.1504/IJPP.2006.010843
Williams, D. W. (2017). Preparing data for forecasting. In J. Sung & TD. Lynch (Eds.), *Government budget forecasting: Theory and practice* (pp. 365-396). Routledge.
Williams, D. W., & Calabrese, T. D. (2016). The status of budget forecasting. *Journal of Public and Nonprofit Affairs*, *2*(2), 127-160. https://doi.org/10.20899/jpna.2.2.127-160
Williams, D. W., & Kavanagh, S. C. (2016). Local government revenue forecasting methods: Competition and comparison. *Journal of Public Budgeting, Accounting & Financial Management*, *28*(4), 488-526. https://doi.org/10.1108/JPBAFM-28-04-2016-B004
Williams, D. W., & Miller, D. (1999). Level-adjusted exponential smoothing for modeling planned discontinuities. *International Journal of Forecasting*, *15*(3), 273-289. https://doi.org/10.1016/S0169-2070(98)00083-1
Wong, J. D. (1995). Local government revenue forecasting: Using regression and econometric revenue forecasting in a medium-sized city. *Journal of Public Budgeting, Accounting & Financial Management*, 7(3), 315-335. https://doi.org/10.1108/JPBAFM-07-03-1995-B001

Zhang, G. P., & Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, *160*(2), 501-514. https://doi.org/10.1016/j.ejor.2003.08.037

Zhang, W., Cao, Q., & Schniederjans, M. J. (2004). Neural network earnings per share forecasting models: A comparative analysis of alternative methods. *Decision Sciences*, *35*(2), 205-237. https://doi.org/10.1111/j.00117315.2004.02674.x