Journal of Business Research 68 (2015) 1810–1818

Contents lists available at ScienceDirect

Journal of Business Research

Improving forecasts for noisy geographic time series ☆

Samuel H. Huddleston [a,*], John H. Porter [b], Donald E. Brown [a]

[a] Department of Systems and Information Engineering, 151 Engineer's Way, P.O. Box 400747, Charlottesville, VA 22904, United States
[b] Department of Environmental Sciences, Clark Hall, University of Virginia, Charlottesville, VA 22903, United States

ARTICLE INFO

Article history:
Received 1 April 2013
Received in revised form 8 July 2013
Accepted 14 December 2013
Available online 22 April 2015

Keywords:
Holt–Winters
ARIMA
Top-down forecasting
Crime forecasting
Geographic time series

ABSTRACT

This study compares the performance of several simple top-down forecasting methods for forecasting noisy geographic time series to the performance of the three methods most commonly used for this problem: naive methods, Holt–Winters (exponential) smoothing, and the ARIMA (Box–Jenkins) class of models. The problem of producing weekly burglary forecasts at the precinct and patrol sector level in the city of Pittsburgh over a five-year period provides a case study for performance comparison. All top-down forecasting methods improve forecasting performance while significantly reducing the modeling workload. These results suggest that simple top-down forecasting models may provide a general-purpose method for improving forecasting for noisy geographic time series in many applications.

Published by Elsevier Inc.

1. Introduction

Most businesses record some form of geographic data: sales, orders, deliveries, or profits by location (Leonard & Samy, 1998). Many businesses also regularly use time series methods to forecast future demand. However, many businesses rarely leverage available geographic data in developing these demand forecasts, usually relying instead on univariate time series methods or leading indicator models.

In a similar vein, police officials regularly use forecasts of future criminal activity to make resource allocation decisions. Police agencies use forecasts of crime counts to divide limited resources across precincts or car beats, assign or redeploy special details based on forecasted demand, and conduct performance assessment of subordinate units (Cohen & Gorr, 2005; Gorr, Olligshlaeger, & Thompson, 2003). The most used forecasting methods in police applications are the naive approaches of referencing the crime count from the previous time period (i.e., previous week or month) or the crime count from the same period twelve months earlier (Gorr et al., 2003). Police use of the univariate time series methods often employed in business remains relatively rare (Cohen & Gorr, 2005).

Both business and law enforcement stand to benefit by the development of accurate models for forecasting noisy geographic time series. Geographic time series are counts of events (sales, deliveries, or crimes) indexed over time by geographic region. Often, geographic time series are very noisy, with the observed counts by region varying considerably from period to period and region to region. This variance is expected since many geographic time series are the result of the superposition of many low-intensity point processes. Examples of low-intensity point processes include the actions of individual criminals in geography, taxi service requests by individual consumers, or pizza delivery requests by individual households. The Palm–Khintchine theorem asserts that the superposition of many low intensity independent renewal processes behaves asymptotically as a Poisson process (Heyman & Sobel, 1982). Thus, when the actions of many individuals acting independently in geography generate a geographic time series, the time series should exhibit the noisy behavior of a Poisson process, in which the variance of event counts in a period is equal to the average count of a period. These noisy geographic time series are very difficult to forecast accurately.

This study compares the performance of several simple top-down forecasting methods for forecasting noisy geographic time series to the three methods most commonly used to forecast them: naive methods, Holt–Winters smoothing, and the ARIMA (Box–Jenkins) class of models. Geographic top-down forecasting models proportion a domain or aggregate level forecast across domain sub-regions based on a geographic weighting scheme, simplifying the modeling process by requiring only one forecasting model rather than unique models for each geographic sub-region. This study compares modeling performance for the problem of producing regular weekly forecasts at the precinct and patrol sector level for burglaries (a very resource intensive crime type) in the city of Pittsburgh over a five-year period. All top-down forecasting methods improve forecasting performance while significantly reducing the modeling workload. These results suggest that simple top-down

☆ Special thanks to Wilpen Gorr of Carnegie Mellon University for providing the data used for this study and to Jon Fox of the U.S. Army, Kesten Green of the University of South Australia, and the anonymous reviewers for their insightful comments and suggestions, which significantly improved the quality of this study.

* Corresponding author.
E-mail addresses: shh4m@virginia.edu (S.H. Huddleston), jhp7e@virginia.edu (J.H. Porter), deb@virginia.edu (D.E. Brown).

http://dx.doi.org/10.1016/j.jbusres.2015.03.040
0148-2963/Published by Elsevier Inc.

modeling approaches may provide a general purpose method for improving forecasts for noisy geographic time series in many applications.

2. Background

As Hyndman and Khandakar (2008) note, "Automatic forecasts of large numbers of univariate time series are often needed in business." Commonly employed univariate time series methods include: time series regression, moving average models, exponential smoothing models (including extensions for seasonality and trend), Auto-Regressive Moving Average (ARMA) models, Auto-Regressive Integrated Moving Average (ARIMA) models, ARMA models with exogenous inputs (ARMAX), spectral time series models, state-space time series models, and generalized auto-regressive conditionally heteroscedastic (GARCH) models (Shumway & Stoffer, 2006). Of the many available time series methods used, the most commonly used (Hyndman & Khandakar, 2008) are the ARIMA family of time series models developed by Box and Jenkins (1990) and the various exponential smoothing methods developed by Brown (1959, 1963), Holt (1957), Holt, Modigliani, Muth, and Simon (1960), and Winters (1960).

Exponential smoothing and ARIMA models are the most popular time series methods for two reasons. First, both methods provide very flexible modeling frameworks that are robust to many types of time series patterns such as trends, seasonality, or unusual changes in the pattern such as the introduction of shocks (Hyndman & Khandakar, 2008). Second, even non-statisticians can easily automate exponential smoothing (Holt–Winters) and ARIMA forecasting models using widely available software. For example, Microsoft Excel easily optimizes the parameters of a Holt–Winters smoothing model using Solver and freely available statistical software such as R provides algorithms for automatically fitting ARIMA models (Hyndman & Khandakar, 2008). Thus, exponential smoothing and ARIMA methods provide benchmarks for the performance of any new method for recurring short-term demand forecasts in many applications.

2.1. Crime forecasting

Police use crime forecasts to divide limited resources among subordinate commands, deploy special details to the most at-risk areas, assess the performance of police units, and schedule training or vacations during non-peak periods (Gorr et al., 2003). The most used forecasting methods in police applications are the naive methods of referencing the crime count from the previous time period (i.e., previous week or month) or the crime count from the same period twelve months earlier (Gorr et al., 2003). Gorr et al. (2003) demonstrate that the Holt–Winters exponential smoothing approach can significantly improve upon naive forecasting for crime forecasts at the precinct level, and the crime analysis and forecasting literature contains almost exclusively time series models from the nested exponential smoothing (Holt–Winters) family. Exceptions to this rule include applications of ARIMA models used to forecast city-level crime rates (Chamlin, 1988; Chen, Yuan, & Shu, 2008a, 2008b) and several studies incorporating leading indicator models. Leading indicator crime models are designed to "make accurate forecasts of the relatively rare, large changes in crime" (Cohen, Gorr, & Olligschlaeger, 2007). Although leading indicator models out-perform traditional forecasting methods in predicting exceptional conditions, such as spikes in violent crime, they often provide very poor performance in forecasting event counts under normal conditions, making them unsuitable for recurring short-term demand forecasts (Cohen et al., 2007; Gorr, 2009).

2.2. Top-down forecasting

Top-down forecasting uses weights or percentages to divide an aggregate forecast across product lines or geographic regions. Top-down forecasting is the opposite of disaggregated forecasting, in which unique forecasts are made for all of the subordinate product lines or regions. These disaggregated forecasts are often subsequently aggregated to produce an overall forecast, a method often referred to as bottom-up forecasting. Top-down forecasting is highly debated in many disciplines, including sales forecasting, product demand forecasting, and econometrics (Widiarta, Viswanathan, & Piplani, 2007). Many empirical studies have explored top-down forecasting, with some studies advocating for the top-down forecasting approach (Ballou, 1999; DeLurgio, 1998; Fogarty, Hoffman, & Blackstone, 1991; Gross & Sohl, 1990; Kahn, 1998) and other studies asserting that disaggregated forecasts should provide similar or better results (Dangerfield & Morris, 1992; Diebold, 1998; Dunn, Williams, & Spivey, 1971; Gordon, Morris, & Dangerfield, 1997; Shlifer & Wolff, 1979). The empirical results in many of these studies are largely unexplained (Birmingham & D'Agostino, 2011). As a result, the general rule of thumb is to comprehensively fit both top-down and bottom-up forecasting models and use the results of each method to better inform the other models (Allen & Fildes, 2001; Lapide, 2006). This approach requires fitting and analyzing a very large number of models in practical application. A simulation study (Huddleston & Brown, 2013) demonstrates that when trend or seasonality effects apply equally to all geographic sub-regions within a domain, a top-down forecasting approach should significantly improve performance. Appendix A provides a brief statistical motivation for why top-down weighted forecasts should improve performance in these cases. This study examines whether the results observed in the simulation study apply in an operational setting.

3. Problem definition

To formally define the forecasting problem considered in this study, let Y_t denote the number of events that occur within the domain of interest D during the time period t. The domain of interest D contains regions indexed by j: $\{D_1, D_2,\ldots, D_J\}$. The quantity of interest is Y_{jt}: the event count for each of the J regions during each time period t. The problem considered is the regular production of one-step ahead forecasts for the noisy geographic time series indexed by Y_{jt}.

4. Data

This study examines the real-world problem of forecasting weekly burglary counts (a very resource-intensive crime type) over a nearly five year period for the City of Pittsburgh. Fig. 1 provides a time series of weekly burglary counts at the city level from 1 January 2006 through 31 October 2010. Fig. 2 provides an illustration of the location of burglaries over a one-year period and the geographic precincts and car beats that define the regions for the forecasts. Gorr and Kurland (2012) provide the geographic data for the City of Pittsburgh in a recently published GIS tutorial designed to "teach crime mapping and analysis skills using ArcGIS Desktop software." As the authors note, "… this book uses real crime data obtained from the Pittsburgh Police Bureau and the Allegheny County 911 Center in Pennsylvania." As Fig. 2 illustrates, the city of Pittsburgh uses 46 patrol car sectors that are made up of an average of three census tracts each. These patrol units are overseen by the six police precincts, with each precinct containing seven to nine patrol car sectors. The forecasting problem is the weekly requirement to generate one-week-ahead burglary forecasts for the 46 police sectors and six precincts. This problem requires a total of 52 one-step-ahead forecasts on a weekly basis.

5. Methodology

This study uses a rolling horizon design to measure out-of-sample forecast accuracy (Makridakis, Wheelwright, & McGee, 1983), which has been used in previous crime forecasting studies in Pittsburgh