CS-BIGS 1(1): 47-53 http://www.bentley.edu/csbigs/vol1-1/holcomb.pdf

Forecasting Police Calls during Peak Times for the City of Cleveland

John P. Holcomb, Jr.
Cleveland State University, USA

Norean Radke Sharpe
Babson College, USA [1]

For a period of time, professors from the Cleveland State University worked closely with the City of Cleveland Police Department. This partnership resulted in access to police records cataloging all emergency 911 calls for the city since 1995. Here, we describe forecasting approaches that can be used by the Police Department based on hourly 911 calls in the years 2001 to 2003 throughout the city during peak call time: the third shift during summer months. This case study is appropriate for class discussions in advanced courses in statistics to explore the application of time series analysis techniques.

Introduction

The city of Cleveland, Ohio is a metropolitan area located on the southern shores of Lake Erie. At the time of the 2000 census, the city had 478,403 residents and the city population had been declining steadily for the last fifty years from a 1950 high of 914,808 residents. With dwindling tax revenues, the city government prioritized maximizing services to residents without increasing costs. In addition, the city was implementing a data-based approach to monitoring the dispensing of city services.

At this same time, Cleveland State University, a comprehensive and government supported institution of higher education located in downtown Cleveland, began to offer financial rewards to faculty who reached out to community agencies to form research partnerships. Faculty in the mathematics and sociology departments began a dialogue with the Cleveland City Police Department to collect and analyze data related to calls of service. At that time, there was no comprehensive analysis of the volume, location, and nature of the calls.

As a result of this partnership, we received access to approximately 5 million records for calls for service (commonly known in the United States as 911 calls) and officer initiated police activity from 1995-2003. A summary of findings from these data appears in Batizy (2004). This paper reports findings on a subset of the 1,721,576 calls made from January 1, 2001 through December 31, 2003.

[1] The authors acknowledge the assistance of Babson College students Tim Consilvio, Mark Dayvie, and Michael Perkins in cleaning data and exploring forecasting models.

The goal of this paper is to describe effective forecasting models for police activity in the city of Cleveland to assist the Cleveland Police Department with staffing during peak crime hours. Specifically, we focus on the peak shift and the peak summer months. Inaccurate forecasts of calls for service are costly to police departments because they result in either an inadequate number of police to respond to the calls, or an oversupply of police on duty. Understaffing of police increases the possibility of delayed time of response to the 911 calls and the associated consequences of a poor response time. Overstaffing increases the direct labor cost to the City of Cleveland.

In addition to the goal of creating an accurate forecasting model, we want to develop a relatively simple user-friendly model for implementation at the Cleveland Police Department. The resulting model needs to be feasible, both from a financial and practical perspective. The reason we chose the peak months and the peak shift was the fact that peak crime times and months present the most challenges in staffing across the districts in Cleveland.

Background

Research in the field of emergency management and response can be found in the operations research, criminology, and sociology literatures. Goldberg (2004) provides a review of operations research and statistical models for the deployment of emergency service vehicles. Goldberg highlights the study of Kamenetzky, Shuman, and Wolfe (1982) that developed a regression model to predict demand for pre-hospital care. Its limitations were that it ignored the temporal issues known to affect demand for services. Mabert (1985) developed a demand-for-911 service model using Box-Jenkins time series methodologies. Recently, Channouf et al. (2007) developed time-series models of call volume for emergency medical service in Calgary and Alberta and Taylor (2007) provided a comparison of five univariate time series methods for forecasting intraday arrivals at a call center in the United Kingdom.

It is widely known that ARIMA models are useful in modeling call center data. Examples include Bianchi, et al. (1998), Andrews and Cunningham (1995), and Nijdam (1990). Most recently, Burman and Shumway (2006) used the ARIMA method to model a U.S. energy and global temperature time series to compare alternative modeling approaches.

In the criminal behavior literature, temporal variations have been documented frequently. Harries, et al. (1984) reviewed the long-associated link between weather and human behavior and provided an analysis of seasonality and assault in Dallas, Texas. Other research relating crime and seasonality includes Anderson, et al. (1997); Auliciems and DiBartolo (1995); Baumer and Wright (1996), Cohn (1993); Cohn and Rotton (2000, 2005); Harries (1990); and Rotton and Cohn (2000, 2004). In many of these studies, it was not uncommon to use data sets that contained crime data from the past decade (or even longer) due to the lag time in reporting and releasing data that are reliable.

Data Description

The first step in our data preparation was aggregating calls to the number of calls per hour. This provided 24,000 data points across all kinds of calls for service and officer calls of police activity. For example, a summary of the Cleveland data shows that there are 219 different identifying codes for the types of calls for police service. The table below gives the top ten categories for calls for service and officer activity.

In Table 1, perhaps surprisingly, the third most common type of calls for service is domestic violence involving assault or threats of assault with the suspect on the scene. Another area of surprise may be the large number of silent 911 calls. These calls are made to the call center, but the caller immediately hangs up or is silent. The City of Cleveland has a policy in place that officers respond to the geographic location of the call to determine if there is need of police service, so each of these requires a response.

Table 1. Frequency of top ten categories shows distribution of types of calls

Nature of Call	Count	Percent
Traffic Stop	186,123	10.8
District Assignment	79,032	4.6
Domestic Violence (Suspect present)	77,420	4.5
Burglar Alarm	71,686	4.2
Silent 911 Call	71,620	4.2
Residential Burglar Alarm	62,889	3.7
Drug Activity	54,195	3.1
Civil Disturbance	50,087	2.9
Suspicious Activity	34,750	2.0
Theft Report	34,312	2.0

Of interest to city council members is the number of calls per service broken down by District. The city of Cleveland has six districts with some known for higher crime than others. A simple frequency table for the six districts of Cleveland reveals a fairly even distribution of calls across the six districts (Table 2).

If, however, we take into account the population of each district, we see a different picture. The per capita analysis shows that district three has the greatest number of calls per capita because, although it has the fewest number of calls, it also has the smallest population of all the districts. (This zone contains the downtown Cleveland business district which contains many businesses, restaurants, and night clubs and fewer residential areas.) While it is interesting to note the differences in calls per capita across the districts, each call still requires a response, and Table 2 demonstrates that the districts are similar in frequency of calls.

Table 2. Frequency of 911 calls by district shows relative consistency across district

District	Calls	Percent
D1	290099	16.9
D2	296278	17.2
D3	270717	15.7
D4	309119	18.0
D5	271219	15.8
D6	284144	16.5

Another variable of interest is the Priority of the service call. The City of Cleveland classifies calls into five priorities with priority 1 calls being the most important. Priority 1 calls are calls for service where crime is in progress such as homicide, domestic violence, ethnic intimidation, shots being fired, or robbery. The distinguishing feature is that the perpetrator is on the premises or close to the premises. Priority 5 calls are often police - initiated calls to the dispatch center that refer to patrolling one's district, filing reports, or walking through public areas such as parks and neighborhoods.

If we examine the number of calls over time by Priority, we observe that seasonal patterns are fairly pronounced across priorities 1, 2, and 3, while priority 4 and 5 calls are more consistent (see Figure 1). Seasonal peaks in the summer are clearly visible in all districts for the most urgent calls: priorities 1, 2, and 3. Each of these time series is stationary in the mean with seasonal peaks during the summer months (June, July, and August) during the years 2001 – 2003.

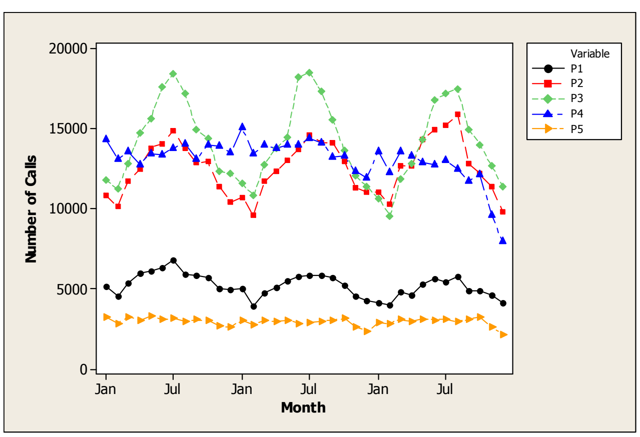

Figure 1. Time series of monthly calls by Priority (P1 – P5) during 2001 – 2003

Since we have hourly data, it is also useful to examine a time series plot of the mean number of calls per hour of the day aggregated across all districts and priorities (Figure 2). Here we see that the busiest hours of the day for calls for service are in the late afternoon/ early evening (hours 15 – 20). This is when traffic is at its worst and young people are home from school. Many individuals believe that police activity is busiest at night, but we see there is a steady decline in calls for police activity from 9:00 pm until 6:00 a.m.

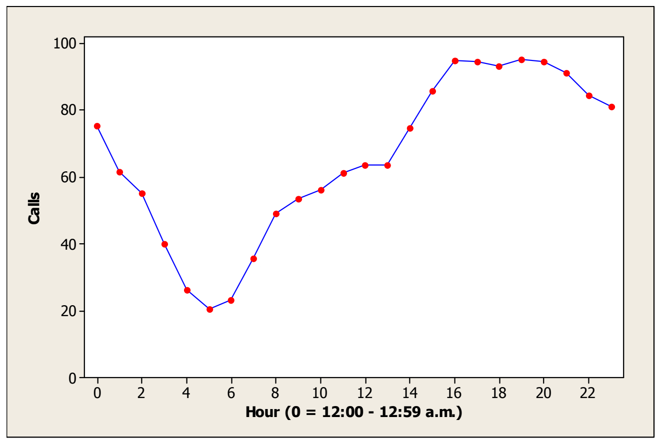

Figure 2. A plot of average calls per hour over the 24 hours in a day shows the peak in calls between hours 15 and 20 (3pm – 8:59pm)

If we examine this six hour time frame (3:00 pm to 8:59pm) across the year on a monthly basis, we obtain Figure 3. Here we see that calls for service do peak during the hot summer months. The median number of calls in June (month 6) is 102 while in December (month 12), the median is 74. Thus, a 38% increase in median number of calls exists comparing June to December. This should not be surprising, since prior studies have shown that summer months associated with warmer weather are likely to be months with more violence (Anderson et al., 1997).

The Cleveland City Police Department schedules police in three different shifts and these peak hours of incoming calls overlap most with the third shift (3PM – 11PM). Thus for staffing purposes, the number of hourly calls will be forecasted during shift three in the summer months of June, July, and August – the peak hours and peak months.

Beside the hourly calls, other variables in our dataset included: seasonal dummy variables, including the day of the week and hour of the day and weather-related variables, such as wind speed, sea level pressure, and temperature (in °F). We will compare models using two different time series. First, we will use all hourly calls during shift 3 between Jan., 2001 and Aug, 2003 (7671 data points). Second, to explore the ability in the future to use smaller datasets, we will use only the hourly calls during peak hours and peak summer months in 2001 – 2003 (2096 data points). In both time series, we held out the observations from August 18, 2003 to August 31, 2003 (112 points, or 14 days of shift 3) to evaluate and compare the accuracy of our models.

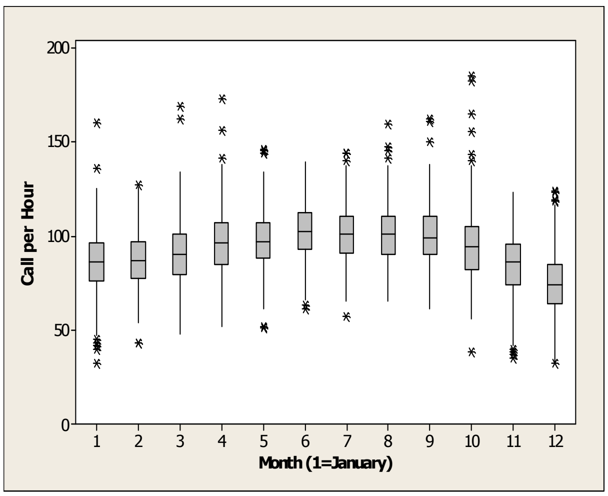

Figure 3. A plot of calls per hour between 3:00pm – 8:59pm by month during 2001 – 2003 shows a summer peak during June – August (month 6 – 8).

Data Analysis

While a large number of models were created, this case study examines the ones we consider the most useful for the Cleveland Police Department. The main metrics used for model evaluation were retrospective and prospective mean absolute percent error (MAPE). After examining the stationary time series and exploring the data, we decided to start with smoothing models. Of these models, the most accurate was an eight-hour moving average model to predict hourly call volume for two weeks in the future – in this case the last two weeks in August, 2003. The reason for the effectiveness of this model is largely because the data were only for one shift so the data will most likely show an eight hour period. As we can see in Figure 4, the 8-hour moving average model showed good results with a relatively low MAPE of approximately 13% using the shorter time series

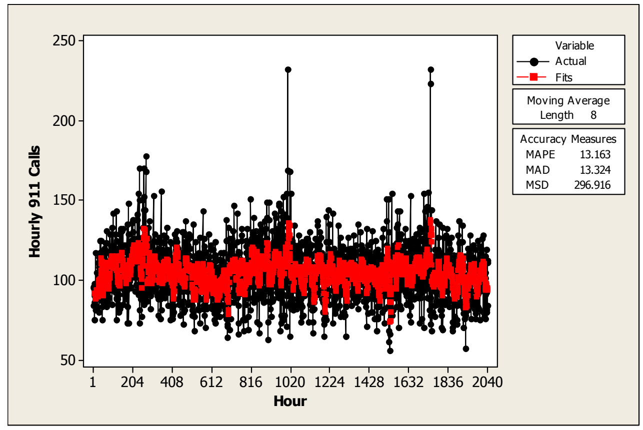

Figure 4. Plot of hourly 911 calls and moving average model during shift 3 for summer months 2001 - 2003 shows a stationary time series with relatively few unusual observations.

containing only prior summer months. (The MAPE was similar at 14% for the lengthier time series).

Note that the moving average model follows the peaks and valleys of this stationary series somewhat closely. The unusually high call observations occurred on July 4th, 2002 and 2003 and on August 14, 2003. The outliers on July 4th are not unexpected, since holidays are known to be likely dummy variables in a crime or sociological regression model. The outlier on August 14, 2003 is most likely due to the onset of a citywide power outage, which was part of a larger blackout and included all of Ohio, the Northeastern United States, and Eastern Canada.

While the FPE (Forecast Percent Error) of this smoothing model was only 3.06% for the last two weeks in August, this simplistic model may not be practical for forecasting two weeks ahead; although it may be accurate on average, it may also miss any volatility during the forecasted two week period. A second smoothing model that was tried was the single exponential model, which produced a similar MAPE (13% and 14% for the short and lengthier time series, respectively).

Given the seasonality of the time series, as a next step in model building, we evaluated the Holt-Winters multiplicative model to capture the level, trend, and seasonal components in the 911 call time series. While the MAPE for the Holt-Winters model was lower (12% and 13%, for the two time series, respectively), the forecast percent error for the last two weeks in August was higher (-11 % for both models) than the simpler smoothing model, which meant that the number of incoming 911 calls were under predicted on average. In fact, approximately two-thirds of the hourly forecasts for the third shift during these last two weeks in August were under forecasts. The implication is that the police force would be consistently understaffed during the third shift if this model were used.

Next, we chose to focus our efforts on a multiple regression approach using temperature, intervention variables, and autoregressive terms. The final model using the shorter time series included dummy variables for the July 4th holiday, day of the week, hour of the day (all with $p < 0.01$) – and quantitative variables for temperature (in °F, $p < 0.01$) and the number of 911 calls 14 and 21 days ago (or 112 and 168 hours ago, $p < 0.05$). Note that more recent autoregressive terms could not be used in the model if the purpose is to forecast calls two weeks into the future. Also, as expected the coefficients support that calls are highest on Fridays between the hours of 9 – 11 PM.. The multiple regression model for the lengthier time series did not include temperature, since it was missing for a period during 2002, but did include dummy variables for the months, in addition to hour and day of the week (all coefficient p-values < 0.001, see Table 3).

Despite our use of dummy variables and autoregressive terms, this model continued to have difficulty with the previously identified outliers and autocorrelation remained a problem. However, prospective error (MAPE of the forecasts) for both time series was the best of all the models. If forecasted temperatures can be reliably obtained two weeks out to predict staffing needs two weeks in advance, then the first model can be used to effectively forecast 911 calls. Otherwise, staffing could be determined from using only the dummy variables for hour, day of the week, and month, plus the level of calls 2 and 3 weeks ago based on the lengthier time series.

Finally, our analysis shifted to autoregressive/ integrated/ moving average (ARIMA) models, originally described by Box and Jenkins (1976). These models use autoregressive (AR) terms, which are lagged observations of the dependent variable and moving average (MA) terms, which are lagged error terms as independent, or explanatory, variables. If the dependent time series is differenced, then the AR and MA terms are integrated (I) with the process of transformation and the predicted variable is now the change, or difference, in the dependent variable.

Our first step in developing an ARIMA was to examine the autocorrelation function (ACF) and the partial autocorrelation function (PACF) to determine the appropriate AR and MA terms to include in the model. The ACF of hourly 911 calls clearly showed the seasonality in the series (see Figure 5) and the PACF indicated additional AR and MA terms that were appropriate for the model. The strong seasonality suggested a seasonal difference, so we used these differenced values as our dependent variable. The resulting ARIMA model included AR terms for the prior hour and the same hour of the day for the five previous days (5 seasonal AR terms). None of the MA terms proved to be significant. Thus the final ARIMA model was (1,0,0)(5,1,0)8. This model produced an improved MAPE over the Holt-Winters method (approximately 12% using both time series). Table 3 compares the error for each of the models discussed in this case study.

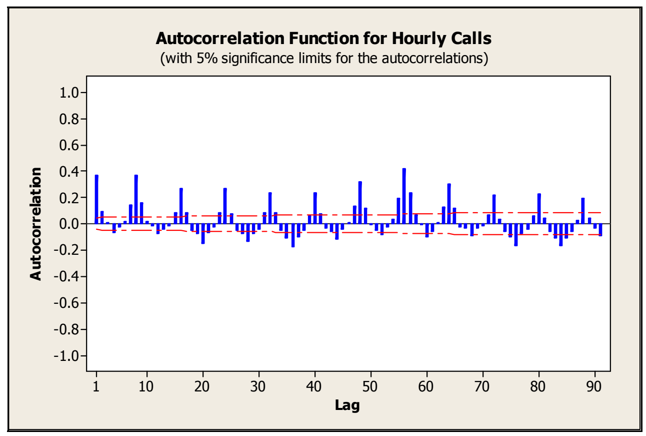

Figure 5. The autocorrelation function shows the relationship between incoming 911 calls to the number of calls the same time on prior days (using only the 8 hours in shift 3).

Conclusion

Our analysis supported previous statistical studies that found a relationship between peak crime rates and temperature. A multiple regression model using temperature (in oF) and intervention variables for holiday, day of the week, and hour of the day together yielded an historical MAPE of less than 10%. In addition, our analysis revealed little difference between models developed on the full time series of hourly calls during shift 3 for every month from Jan., 2001 to Aug., 2003 compared to the shorter time series (using only prior summers). The model developed on the lengthier time series did, however, produce the lowest prospective MAPE and enables the Cleveland Police Department to forecast calls based only on the hour, day, month, July 4th holiday, and the level of calls from 2 and 3 weeks ago. The advantage of this approach is that the coefficients from this model allow police to build a scheduling model

Table 3. Comparison of error for the models.1

	Using Shorter Time Series2		Using Lengthier Time Series3	
	Retrospective MAPE	Prospective MAPE (for forecasts)	Retrospective MAPE	Prospective MAPE (for forecasts)
Multiple Regression4	9.28%	13.72%	12.63%	10.93%
Holt-Winters Method	12.01%	17.68%	13.40%	17.68%
ARIMA (1,0,0)(5,1,0)8	11.08%	16.31%	12.70%	16.19%

1. Theil's <1 for each model, indicating that the model is more accurate than a naïve forecast.
2. Using Hourly Calls during June-Aug in 2001 – 2003 for shift 3
3. Using Hourly Calls during all months in 2001 – 2003 for shift 3
4. Multiple Regression Model for lengthier time series w/ lowest prospective MAPE:

Calls = 43.0 – 6.85Sun + 2.01Mon + 4.55Tue + 4.75Wed + 3.18Thu + 5.38Fri + 4.28Hr15 + 6.06Hr16 + 4.73Hr17 + 6.19Hr18 + 5.82Hr19 + 4.53Hr20 + 1.57Hr21 + 19.2Holiday + 0.171Lag112 + 0.177Lag168 + 5.15Jan + 5.29Feb + 7.82Mar + 14.8Apr + 13.8May + 23.1Jun + 17.8July +17.8Aug + 15.7Sep + 9.49Oct + 4.26Nov

two weeks in advance using variables that are easily accessible.

The need to have a model that is feasible and flexible supports the use of models that can be easily explained, programmed, and implemented. The hardware and software requirements of our models are minimal and the forecasts can be readily obtained and updated. As with any forecasting models, the uncertainty of events, such as major storms or blackouts, contribute to the volatility of the series and of the forecasting error. In addition, limitations to model implementation include labor union rules, advance posting of shifts, and the desire of many officers for vacation time during the summer season; these all provide challenges for the city to maximize resources in the peak summer months. These models do provide guidance, however, on the expected staffing needs during the busiest period of calls for service to the City of Cleveland.

References

Anderson, C. A., B.J. Bushman, and R.W. Groom. 1997. Hot years and serious and deadly assault: Empirical tests of heat hypothesis, *Journal of Personality and Social Psychology*, 73(6): 1213-1223.

Andrews, B. H. and S.M. Cunningham. 1995. L. L. Bean Improves Call Center Forecasting, *Interfaces*, 25(6): 1-13.

Auliciems, A. and L. DiBartolo. 1995. Domestic violence in a subtropical environment: Police calls and weather in Brisbane, *International Journal of Biometeorology*, 39: 34 - 39.

Batizy, L. 2004. Using quantitative methods to model urban police: Calls for service, *STATS*, 40: 14-18.

Baumer, E. and R. Wright. 1996. Crime seasonality and serious scholarship: A comment on Farrell and Pease, *British Journal of Criminology*, 36(4): 579-581.

Bianchi L., J. Jarrett and C.R. Hanumara. 1998. Improving forecasting for telemarketing centers by ARIMA modeling with intervention, *International Journal of Forecasting*, 14(4): 497-504.

Box, G. E. P., and G. M. Jenkins. 1976. *Time Series Models for Forecasting and Control*, Holden Day, San Francisco, CA.

Burman P. and R. Shumway. 2006. Generalized Exponential Predictors for Time Series Forecasting, *Journal of the American Statistical Association*, 101(476): 1598-1606.

Channouf, N., P. L'Ecuyer, A. Ingolfsson and A. N. Avramidis. 2007. The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta, *Health Care Management Science*, to appear.

Cohn, E. G. 1993. The prediction of police calls for service: The influence of weather and temporal variables on rape and domestic violence, *Journal of Environmental Psychology*, 13: 71-83.

Cohn, E.G. and J. Rotton. 2000. Weather, seasonal trends and property crimes in Minneapolis, 1987-1988: A moderator-variable time series analysis of routine activities, *Journal of Environmental Psychology*, 20: 257-272.

Cohn, E. G., and J. Rotton. 2005. The curve is still out there: A reply to Bushman, Wang, and Anderson's 2005 Is the curve relating temperature to aggression linear or curvilinear? *Journal of Personality and Social Psychology*, 89: 67-70.

Goldberg, J. 2004. Operations Research Models for the Deployment of Emergency Services Vehicles, *EMS Management Journal*, 1(1): 20-39.

Harries, K. D. 1990. *Serious violence: Patterns of homicide and assault in America*, Springfield, IL, Charles C. Thomas.

Harries, K. D., S. J. Staler and R. T. Zdorkowski. 1984. Seasonality and assault: Explorations in inter-neighborhood variation, Dallas 1980, *Annals of the Association of American Geographers*, 74(4): 590-604.

Kamenetzky, R., L. Shuman, and H. Wolfe. 1982. Estimating need and demand for prehospital care, *Operations Research*, 30(6): 1148-1167.

Mabert, V. 1985. Short interval forecasting of emergency phone call (911) work loads, *Journal of Operations Management*, 5(3): 259-271.

Nijdam, J. 1990. Forecasting telecommunications services using Box-Jenkins (ARIMA) models, *Telecommunication Journal of Australia*, 40(1): 31-37.

Rotton, J. and E. G. Cohn. 2000. Weather, disorderly conduct, and assaults: From social contact to social avoidance, *Environment and Behavior*, 32(5): 651-673.

Rotton J., and E. G. Cohn E. G. 2004. Outdoor temperature, climate control, and criminal assault: the spatial and temporal ecology of violence, *Environment and Behavior*, 26(2): 276-306.

Taylor, J. W. 2007. A comparison of univariate time series methods for forecasting intraday arrivals at a call center. *Management Science*, forthcoming.

Correspondence: sharpen@babson.edu