# The Impact of Special Days in Call Arrivals Forecasting: A Neural Network Approach to Modelling Special Days

Devon Barrow[a,*], Nikolaos Kourentzes[b]

[a]Faculty of Business, Environment and Society,
Coventry University, Coventry, West Midlands, CV1 5FB, UK
[b]Lancaster University Management School
Department of Management Science, Lancaster, LA1 4YX, UK

## Abstract

A key challenge for call centres remains the forecasting of high frequency call arrivals collected in hourly or shorter time buckets. In addition to the complex intraday, intraweek and intrayear seasonal cycles, call arrival data typically contain a large number of anomalous days, driven by the occurrence of holidays, special events, promotional activities and system failures. This study evaluates the use of a variety of univariate time series forecasting methods for forecasting intraday call arrivals in the presence of such outliers. Apart from established statistical methods we consider artificial neural networks (ANNs). Based on the modelling flexibility of the latter we introduce and evaluate different methods to encode the outlying periods. Using intraday arrival series from a call centre operated by one of Europe's leading entertainment companies, we provide new insights on the impact of outliers on the performance of established forecasting methods. Results show that ANNs forecast call centre data accurately, and are capable of modelling complex outliers using relatively simple outlier modelling approaches. We argue that the relative complexity of ANNs over standard statistical models is offset by the simplicity of coding multiple and unknown effects during outlying periods with ease.

*Keywords:* Time Series Forecasting, Call Centre Arrivals, Outliers,

*Correspondance: D.K. Barrow, School of Strategy and Leadership, Faculty of Business and Law, Coventry University, Coventry, West Midlands, CV1 5FB, UK. Tel.: +44-024-7765-7413

*Email address:* devon.barrow@coventry.ac.uk (Devon Barrow)

## 1. Introduction

A key challenge in many call centres is the accurate forecasting of inbound call volumes required to support short, medium and long-term decisions. With 60–80% of the overall operating budget of a call centre resulting from staffing costs (Brown et al., 2005; Aksin et al., 2007), accurate and robust forecasting for workload management is an important issue. In recent years consumer demand for call centre services has increased along with the dramatic expansion in the service industry (Shen and Huang, 2005; Aksin et al., 2007). With this increase in the number of call centres and the amount of call centre data, new challenges related to the handling and forecasting of such data are presented. Firstly, call centre arrivals data are typically high-dimensional and sampled at a high frequency, usually collected in daily or smaller time buckets (e.g. 15 or 30 minute periods). Under these conditions conventional statistical models designed for low frequency time series may break down and be inappropriate (Kourentzes and Crone, 2010). Secondly, call centre arrivals data exhibit complex seasonal patterns (De Livera et al., 2011). The data often contains intraday, intraweek and even intrayear dependencies, meaning that call arrival volumes will typically exhibit multiple seasonal cycles which need to be modelled (for example, see Taylor, 2008a).

Call centre arrival data is also context sensitive, meaning that the data can contain strong effects from holidays, special events, promotional activities and unexplained variations (Andrews and Cunningham, 1995). These will usually exhibit very different arrival characteristics from regular patterns of call arrivals, weakening or even destroying the correlation structures in the data. This may affect model specification and parameter estimation (Chatfield, 2013). This is distinct from the behaviour of electricity load data known to exhibit similar time series seasonal structure. Hence modelling call centre data may require data cleansing. When outliers are due to holiday and promotional effects, these can often be captured through past experience and reference to market information. On the other hand, when outliers are due to systems errors and other unexplained events they become more difficult to identify and model.

Many forecasting models assume as a standard approach that this information is either available or that the forecaster has an external methodology

for tackling outliers. Some methods developed specifically for call centre data acknowledge this issue. For example, Taylor et al. (2006) take the approach of removing such days, while Shen and Huang (2005) describe the application of singular vector decomposition for outlier detection but provide no empirical evaluation. Many others however avoid this problem by assuming the data pre-cleansed (for examples see: Jongbloed and Koole, 2001; Avramidis et al., 2004; Taylor, 2008a; Pacheco et al., 2009; Taylor, 2010b). Conejo et al. (2005) in forecasting electricity prices try to automatically correct outliers using conventional time series modelling approaches, but have limited success as the high frequency nature of the data makes such approaches inappropriate. Nonetheless, in the context of electricity load forecasting Kourentzes (2011) demonstrates that there are substantial accuracy benefits to be had from modelling irregular load patterns.

This study has two aims. First we evaluate a variety of univariate time series forecasting methods that are in theory capable of modelling data exhibiting characteristics of intraday call centre arrivals. This evaluation is central to identifying which method is most fit for purpose (Petropoulos et al., 2014). We identify empirically artificial neural networks (ANNs) as having the best forecasting performance and argue that this is due to their modelling flexibility. This is a useful finding given the limited research in ANNs for forecasting call centre data, despite evidence which shows them capable of handling the complex seasonal structure of this data (Lee Willis and Northcote-Green, 1983; Temraz et al., 1997). We augment the best performing method to model special events and outlying days. We introduce and evaluate multiple alternative methodologies, ranging from including the outlier information in the forecasts to cleaning the data prior to fitting the forecasting model. This addresses a gap in research of practical significance, considering the difficulty and cost associated with manual exploration and treatment of high frequency data by experts, and limited advancements in automatic outlier identification for such data. Note that we restrict our attention to time series methods that forecast call arrivals volume and we do not investigate models for call arrivals rate, for example by consider other classes of methods e.g. those which incorporate stochastic models (Avramidis et al., 2004) and those based on assumptions of a Poisson process (Aldor-Noiman et al., 2009; Ibrahim and L'Ecuyer, 2013).

The rest of the paper is organised as follows: first we describe the nature of call centre data by introducing the dataset for the empirical results. Section 3 presents the evaluation of numerous forecasting methods and assesses the

impact of outliers on the performance of these methods. Subsequently in Section 4 presents and evaluates various alternatives to modelling of outliers using the best performing forecasting method. The final section provides a summary and concluding comments.

## 2. Call centre arrival data and outliers

### 2.1. Case study dataset

To illustrate the data properties of call arrival time series we describe the two time series that are used in the empirical evaluations. These record calls received by a large call centre of one of Europes leading entertainment companies representing inbound sales calls.

The data is sampled at half-hourly intervals and covers a period of 103 weeks and 3 days from 29 June 2012 to 23 June 2014 inclusive, including bank holidays and weekends. Fig. 1 provides examples of the two time series. Observe that between each day there are some hours with zero values, which represent the hours that the call centre is not operational. Furthermore, it is evident that there are some days with extraordinary levels of calls, which we refer to as outlying days.
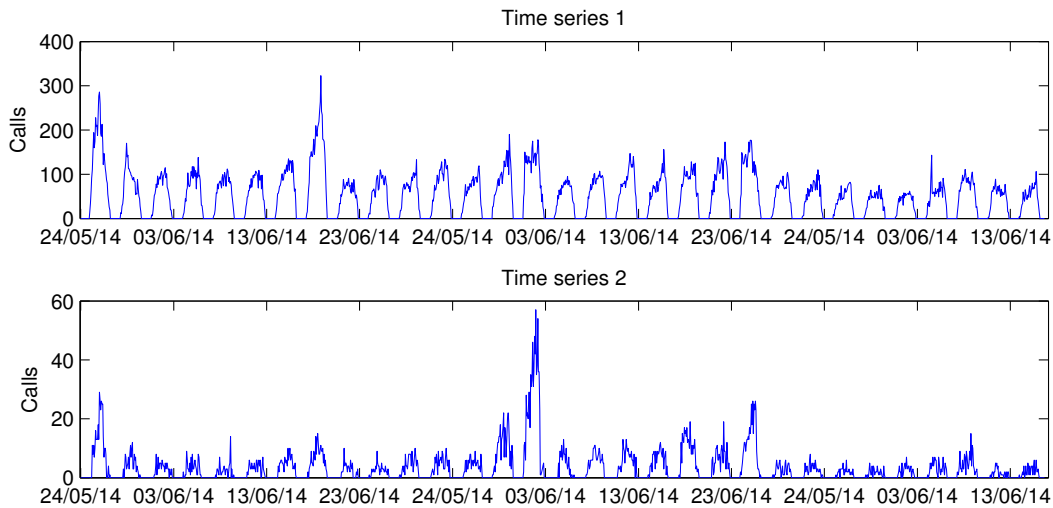


Figure 1: Examples from the two time series from Sunday $25^{th}$ May to Tuesday $24^{th}$ June 2014.

4

Fig. 2 provides a weekly seasonal plot of the first time series. To make the plot less cluttered we provide various percentiles. The difference between the outer bands and the the median provides some insight into the variability of each half-hourly interval of the time series. Two seasonal cycles are evident: an intraday and an intraweek. Monday to Friday look relatively similar, although Monday, Wednesday and Friday exhibit slightly increased afternoon/evening variability. The weekend days are substantially different. Also note that the starting time for each day is not always at 08:00, reflecting slight variations in the starting time that the call centre services calls. The second series exhibits similar seasonal properties.
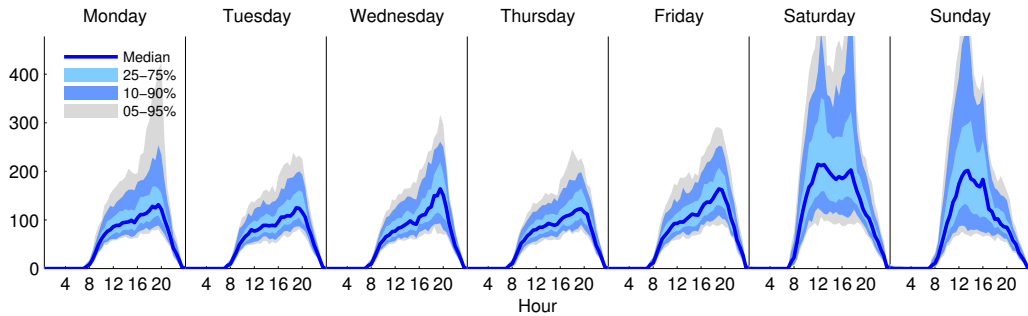


Figure 2: Seasonal percentiles plot for time series 1.

For these time series we are provided with periods of extraordinary activity. These were identified by experts in the organisation that the data originate from, using domain knowledge. In total there are 87 such outlying days for the first time series and 115 for the second.

## 2.2. Outliers in call arrival data

In addition to the regular patterns present in call centre time series a number of irregular patterns exist. One type of irregularity is due to the effects of external factors such as bank holidays and marketing or promotional activities. For example, customer mailings can be geared toward generating increased call centre traffic (Andrews and Cunningham, 1995). These types of shocks can be identified and anticipated fairly accurately, given that they are driven either by actions of the company or calendar events. However, on top of these there are events that are unexpected and an organisation has little information about them a-priori. For example such events might be unexpected problems with products, causing an influx of customer support

5

calls, and so on. A further type of irregularity is posed by unexpected variations in the data due to system and data collection errors. These are often unforeseen and difficult to identify and explain without expensive manual intervention, which is often infeasible given the large amounts of data under consideration.

Regardless the source of the irregular variations call centre data, and in general high frequency data, present a special type of outlier, where the extraordinary behaviour may last several consecutive periods, instead of a single period that is common in low frequency time series forecasting. Therefore, while call arrivals may differ within half-hourly or hourly periods, we consider analysing call data as whole days providing information about outlying days as curves varying over time. For example, call centre arrivals can be different throughout the duration of a bank holiday as compared to a working day. Such outliers can be different both in terms of volume of calls but also the shape of the call pattern over the duration a fixed period, in contrast to conventional outliers that cause only differences in the location of the value of a time series. Consequently, we treat all outliers as full days and therefore all half-hour periods on such days as constituting a single (day) outlier. To better illustrate this Fig. 3 plots the median seasonal weekly pattern of the second time series from the case study and known outlying periods. It can be seen that the outliers vary both in shape and location and diverge substantially from the median of the time series evidence that different modelling and operational responses are required for these 'special' days as a whole.
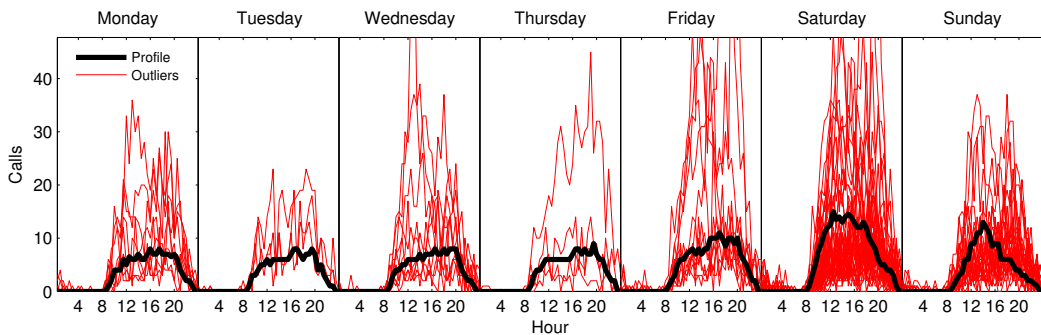


Figure 3: Median seasonal profile and known outliers for time series 2.

There are two challenges that are associated with the presence of outliers in the data. First, they have an impact on the quality of the fitted model;

6

and second, they need to be forecasted in the out-of-sample periods. The presence of these outlying periods can lead to model misspecification and biased parameter estimates (Gelper et al., 2010; Koehler et al., 2012; Chatfield, 2013). Since outliers can last for several periods, typically complete days, these can also bias the estimation of seasonal patterns. This problem is exacerbated for methods that may be overparametrised, such as exponential smoothing with multiple seasonal cycles (Gould et al., 2008), amongst other methods (Draper and John, 1981; Tsay, 1989). When an outlier is expected to occur in the future periods the forecasting methods need to be able to account for this. Given the size and frequency of such outliers in call centre arrival time series it is perhaps not surprising that simple methods like Seasonal Moving Average are found to perform well (Taylor, 2008a).

To the best of our knowledge, the literature on call centre forecasting does not provide guidelines how to deal with such data. Similar forecasting applications, like electricity load forecasting, do not provide clear guidelines either, focusing mostly on forecasting the normal periods.

An alternative is to analyse such outliers as functional data (Ramsay and Silverman, 2002). Instead of considering call arrivals in a traditional time series manner, data can be viewed as curves (for example of daily duration) that are observed sequentially in time (Hyndman and Shang, 2009) and therefore apply functional data analysis (Silverman and Ramsay, 2005). In contrast to normal outliers, functional outliers differ not only in level, but also in shape over the duration of a fixed period. This means that outliers may lie outside the range of the vast majority of data or they may be within the range of the rest of the data but have a very different shape than other curves; the former referred to as magnitude outliers and the latter called shape outliers (Hyndman and Shang, 2010). In some cases outliers will exhibit a combination of both features. Here, in the proposed solutions to modelling outliers in call centre arrivals the notion of function outlier will be useful, as our objective is to track both the differences in shape and magnitude. We therefore consider outliers as whole days across multiple observations as opposed to conventional additive or innovative outliers which consider individual observations. In addition to simplifying modelling complexity by treating each high frequency observation as distinct, this approach provides insights at the aggregate level of days of the week which is valuable for many types of operational and capacity planning decisions.

## 3. Evaluation of call centre forecasting methods

Developments in time series forecasting for call centre arrival data have focused primarily on advanced methods for capturing the complex seasonal patterns in the data as well as dealing with the high-frequency, big data aspects of the problem (for examples see Antipov and Meade, 2002; Shen and Huang, 2005; Taylor, 2008a; Shen, 2009; Aldor-Noiman et al., 2009; Taylor and Snyder, 2012; Ibrahim and L'Ecuyer, 2013). Meanwhile, simple methods such as the intraweek seasonal moving average, have been shown to outperform more advanced ones (Tandberg et al., 1995; Taylor, 2010b; Ibrahim and L'Ecuyer, 2013), with Taylor (2008a) suggesting "to use more advanced methods may not be the solution". On the other hand flexible methods, such as neural networks, have been overlooked, which we attempt to address here.

### 3.1. Forecasting Methods for Call Centre Data
### 3.1.1. Seasonal Naïve

The Naïve method is a fundamental benchmark for forecasting. It can be extended to its seasonal variant as $\hat{y}_{t+h} = y_{t+h-m}$ where $y_t$ is the recorded arrivals at a call centre at period $t = 1, \ldots, n$, $\hat{y}_t$ are the forecasted arrivals, $h$ is the forecast horizon and $m$ the length of the seasonal cycle. As this formulation allows capturing only a single seasonal cycle we evaluate coding both the daily ($m = 48$) and the weekly ($m = 336$) seasonalities separately. Namely using the Seasonal Naïve we define *Naïve Day* and *Naïve Week*.

### 3.1.2. The Seasonal Moving Average

Following the works by Tandberg et al. (1995) and Taylor (2008a, 2010b) the Seasonal Moving Average is considered as a good contender for forecasting call centre arrivals. It is calculated as:

$$\hat{y}_{t+h} = \frac{1}{k} \sum_{i=1}^{k} y_{t+h-mk}, \tag{1}$$

where $k$ is the number of seasonal periods considered in the calculation of the moving average. Similar to the Seasonal Naïve we build for each seasonal length: *MA Day* and *MA Week*. To identify the appropriate order $k$ of the Seasonal Moving Average we measure the sum of squared errors over a validation set and pick the order that results in the minimum error.

### 3.1.3. Exponential Smoothing

Hyndman et al. (2002) embedded the pre-existing exponential smoothing methods within a statistical state space framework: ETS, which is used in this paper. ETS is able to model time series with various characteristics in terms of trend (none, linear, multiplicative, damped or not) and seasonality (none, additive, multiplicative) and how the error term interacts (additive, multiplicative). Combining all different options in terms of trend, season and error up to 30 different model forms of ETS can be described (Hyndman et al., 2008; Ord and Fildes, 2013). The appropriate form can be selected by using information criteria, such as the Akaike Information Criterion (Hyndman et al., 2002), or alternatively combine different forms (Kolassa, 2011; Kourentzes et al., 2014b). In particular for the available call arrivals time series the presence of zero values restricts the model forms to additive and given the seasonality and lack of trend ETS(A,N,A) is a reasonable:

$$
\begin{align}
y_t &= l_{t-1} + s_{t-m}, \tag{2}\\
l_t &= l_{t-1} + \alpha e_t, \tag{3}\\
s_t &= s_{t-m} + \gamma e_t, \tag{4}
\end{align}
$$

where $y_t$, $l_t$ and $s_t$ model the value of the series, the level and seasonality components at time $t$ respectively and the smoothing parameters $\alpha$ and $\gamma$ control how fast each component updates from errors $e_t$. The seasonal period is $m$ can be adjusted accordingly to capture different seasonal cycles. The model requires an initial level and $m$ initial seasonal values. These, together with the smoothing parameters, are optimised using maximum likelihood estimation.

Gardner (2006) observed that for seasonal ETS the number of free parameters increases substantially. This is particularly relevant for high frequency time series, where the number of initial states can become very high. For example fitting a model with no trend and daily seasonal cycle sampled at half-hourly intervals requires 2 smoothing parameters and 49 initial values to be estimated, while for a weekly seasonal cycle the latter increases to 337 initial values. All these would need to be optimised incurring a high computational cost and potential optimisation issues. In this study both seasonal lengths are used, resulting in *ETS Day* and *ETS Week*.

9

### 3.1.4. Double Seasonal Exponential Smoothing

Given the restriction of ETS to modelling a single seasonality research has focused to extend it to model multiple seasonal cycles. Taylor (2003) proposes an extended exponential smoothing method that is capable of capturing two seasonal cycles simultaneously and models error autocorrelation. Taylor (2010a), based on this work, provides an innovations state space model, which is used in this paper and named *ETS Double* hereafter:

$$
\begin{aligned}
y_t &= l_{t-1} + s_{t-m_1} + d_{t-m_2} + \phi z_{t-1} + e_t, & (5) \\
l_t &= l_{t-1} + \alpha e_t, & (6) \\
s_t &= s_{t-m_1} + \gamma e_t, & (7) \\
d_t &= d_{t-m_2} + \delta e_t, & (8) \\
z_t &= y_t - (l_{t-1} + s_{t-m_1} + d_{t-m_2}), & (9)
\end{aligned}
$$

where $d_t$ is the component for the second seasonal cycle. The two seasonal cycles have periods $m_1$ and $m_2$ and the model now requires four parameters to be estimated: $\alpha$, $\gamma$, $\delta$, $\phi$ and $m_1 + m_2 + 2$ initial values. All are optimised using maximum likelihood estimation. Note that Gould et al. (2008) echoing the arguments of Gardner (2006) warns on the implied dimensionality of the initial states of such models.

### 3.1.5. Autoregressive Integrated Moving Average

The Autoregressive Integrated Moving Average (ARIMA) constitute another popular class of models in univariate time series forecasting. While exponential smoothing primarily captures integrated moving average processes, always implying non-stationary time series, ARIMA can capture both autoregressive and moving average processes, and the order of integration permits the modeller to capture either stationary or non-stationary series. ARIMA models can be easily extended to include any number of seasonal cycles by simply multiplying appropriate seasonal polynomials. For a detailed presentation of ARIMA and single seasonal ARIMA the reader is referred to Box et al. (2008). Multiple seasonal ARIMA models are discussed and evaluated in detail by Taylor (2008b), who finds that the Double Seasonal Exponential Smoothing outperforms such models.

Although ARIMA models are very flexible and powerful, at the same time they are very difficult to specify, which has been one of the key reasons for them not being widely applied in practice. This is particularly relevant to

seasonal and multiple seasonal ARIMA for which the potential search space for the polynomials' orders is greatly expanded and an exhaustive evaluation of all alternative specifications quickly becomes infeasible. To overcome this we follow the methodology by Hyndman and Khandakar (2008) who recommend to first identify the appropriate differencing by employing the KPSS and OCSB statistical tests and then proceed to identify the orders of the polynomials using a stepwise search with the Akaike Information Criterion. We build *ARIMA Day*, *ARIMA Week* and *ARIMA Double* in analogy to the exponential smoothing models discussed above.

### 3.1.6. Artificial Neural Network

Artificial Neural Networks (ANNs) in various forms have been successfully applied in multiple forecasting applications (Zhang et al., 1998; Hamid and Iqbal, 2004), demonstrating forecasting performance at least as good as established statistical forecasting methods (Adya and Collopy, 1998). ANNs are well placed to forecast high frequency time series (Kourentzes and Crone, 2010) and in particular seasonal ones, with several examples coming from electricity load forecasting (Hippert et al., 2005), where they are implemented successfully in electric utility operations and standard software packages, such as the ANNSTLF architecture (Khotanzad et al., 1998) employed in over 50 utilities worldwide. Yet there is very limited application of ANNs in call centre forecasting and the results are mixed. Pacheco et al. (2009) apply an Improved Backpropagation Neural Network to forecasting incoming phone calls in a multi-skill call centre. For call groups having complex behaviour the ANN is found to forecast better than exponential smoothing. However for call groups with low incoming flow exponential smoothing is best. Balaguer et al. (2008) predict the number of forthcoming service requests to a support centre using Time Delay Neural Networks, which outperformed ARMA models in predicting the number of requests 6 hours ahead. Other studies apply ANNs but do not compare to any standard benchmarks such as single or double exponential smoothing, making difficult to assess their performance (for examples see Li et al., 2011; Millan-Ruiz and Hidalgo, 2010).

The mixed results can be partially explained by the inconsistent modelling practices in building ANNs. Indeed Armstrong (2006), doing a meta-analysis of empirical evidence in the forecasting literature, argued that ANNs have not advanced sufficiently to permit their widespread use, due to modelling complexity, lack of robustness and transparency, in particular when considered against established statistical methods. However, recent research

has started to address these limitations, revealing modelling best practices (identification of input variables, use of ensembles of networks, etc; see Crone et al., 2011), resulting in reliable forecasts. The most common type of ANNs used for forecasting is the Multilayer Perceptrons (Zhang et al., 1998):

$$\hat{y}_t = \beta_0 + \sum_{h=1}^{H} \beta_h g \left( \gamma_{0i} + \sum_{i=1}^{I} \gamma_{hi} p_i \right), \tag{10}$$

where $I$ denotes the number inputs $p_i$, which can be lags of the target time series or external variables, and $H$ is the number of hidden nodes in the network. The weights $\boldsymbol{w} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$, with $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_H]$ and $\boldsymbol{\gamma} = [\gamma_{11}, \ldots, \gamma_{HI}]$ are for the hidden and output layer respectively. The $\beta_0$ and $\gamma_{0i}$ are the biases of each node, and serve a similar purpose to the intercept in conventional linear regression. The transfer function $g(\cdot)$ may be non-linear and is usually either the sigmoid logistic or the hyperbolic tangent function.

The network described in Eq. (10) has a single hidden layer that has been shown to be sufficient to give the network universal approximation capabilities, when the number of hidden nodes $H$ is sufficient (Hornik, 1991). The adequacy of a single hidden layer is supported by empirical evidence (Zhang et al., 1998). From Eq. (10) it is easy to appreciate the flexibility of MLPs. Any type of input variables can be used to map the behaviour of the time series in hand, linearly or nonlinearly, permitting interactions between the inputs, if any. This flexibility is one of the key strengths of ANNs, but also one of the reasons for the inconsistent and opaque modelling practices in the literature, due to the large number of options.

To produce forecasts with ANNs we need to identify the appropriate input variables, the size of the hidden node and select the network training setup. To select the hidden nodes we employ the methodology described by Crone and Kourentzes (2010) as extended by Kourentzes and Crone (2010), which is based on regression diagnostics to identify the relevant autoregressive inputs for the ANN. The two seasonal cycles of the call arrivals data are coded using pairs of trigonometric dummy variables:

$$s_{1,j} = sin\left(2\pi t / m_j\right), \tag{11}$$
$$s_{2,j} = cos\left(2\pi t / m_j\right), \tag{12}$$

where $m_j = \{48, 336\}$, resulting in four additional inputs. Because of the flexibility of ANNs this parsimonious encoding was found to be adequate

to capture multiple seasonal cycles (Crone and Kourentzes, 2009).The number of hidden nodes is identified experimentally, by tracking the setup that provides the minimum MSE in the validation set.

Standard network training is used, employing Scaled Conjugate Gradient Descent (Møller, 1993), with early stopping by tracking the error on a validation set (Ord and Fildes, 2013). To avoid poor quality solutions the network is trained with different initial weight 50 times and an ensemble of all trained networks is used to produce the final forecast. The median combination operator is used, as it is shown to be more accurate and converge to reliable and stable ANN forecasts with smaller ensembles (Kourentzes et al., 2014a).

*3.2. Experimental design*

Both time series are split into two samples, one used for fitting and one for testing the out-of-sample performance of the various forecasting methods. The latter sample is comprised by the last 100 days (4,800 data points) of the complete sample. The forecast horizon that is used for the evaluation is from 1 to 48-steps ahead, i.e. up to one day ahead forecast. We employ a rolling origin evaluation scheme, collecting for each method 4,753 48-step ahead forecasts, facilitating the comparison between methods. For the Seasonal Moving Average and ANN a validation sample is also required. To this purpose we retain 100 days prior to the test set, when such a sample is needed.

To further validate our results we repeated the experiment by removing the last 100 days of the time series and constructing a new test set with the preceding 100 ones, effectively constructing again 4,753 48-step ahead rolling origin forecasts. Naturally, the new test set included different outlying days. Nonetheless the results from the second test provide the same findings as the first one and therefore for brevity we report and discuss only the detailed results of the first test.

To measure the accuracy of the forecasts the Relative Mean Absolute Error (RMAE) is used (Davydenko and Fildes, 2013):

$$\text{MAE} = \sum_{t=1}^{h} |y_t - \hat{y}_t|, \tag{13}$$

$$\text{RMAE} = \frac{\text{MAE}_{\text{Method 1}}}{\text{MAE}_{\text{Method 2}}} \tag{14}$$

13

We prefer RMAE over other metrics because its calculation is not affected by zero or low values, while being scale independent. If RMAE = 1 then both methods are equally accurate; if RMAE > 1 then Method 2 is more accurate and the opposite is true when RMAE < 1. Here we will be using the MAE of the ANN forecasts as the denominator for all methods.

Finally we track the accuracy in three separate cases: (a) overall; (b) for normal observations; and (c) for outlying observations. This way we can compare the performance of each method under different conditions. The errors are measured across all observations and periods with zero arrivals are not excluded.

### 3.3. Results

The forecast accuracy across both time series for the out-of-sample RMAE performance is summarised in table 1. The best performance in each column is highlighted in boldface.

Table 1: Average RMAE of forecasting methods

| Method | Overall | Outlier | Normal |
|---|---|---|---|
| Naïve Day | 1.416 | 1.165 | 1.529 |
| Naïve Week | 1.383 | 1.107 | 1.500 |
| MA Day | 1.326 | 1.199 | 1.404 |
| MA Week | 1.181 | **0.909** | 1.279 |
| ETS Day | 1.744 | 1.403 | 1.872 |
| ETS Week | 1.504 | 1.227 | 1.608 |
| ETS Double | 1.119 | 0.924 | 1.196 |
| ARIMA Day | 1.254 | 1.180 | 1.288 |
| ARIMA Week | 1.144 | 0.918 | 1.228 |
| ARIMA Double | 1.826 | 1.434 | 1.991 |
| ANN | **1.000** | 1.000 | **1.000** |

The best overall performance is achieved by ANN. This is consistent across normal periods, while for outlying periods MA Week performs best, validating the strong performance of the seasonal moving average observed by Taylor (2008a). Considering the option of capturing a single seasonal cycle, the intraday or intraweek, for all methods the weekly variant performs best consistently across all three cases: overall, outlier and normal. ETS Double substantially outperforms both ETS Day and ETS Week. This is not the case for ARIMA Double, which is outperformed by the simpler ARIMA

Week. This is attributed to the complexity of specifying the orders and optimising the parameters of ARIMA Double.

If ANN are not considered, when comparing across method families we observe that the more complex ETS variants are outperformed by the simpler Naïve and MA. This is explained by the difficulties associated with the specification and parametrisation of such models as discussed in section 3.1.3. Biased parameter estimates which fail to robustly capture seasonal patterns lead to poor performance not only for outlying days but also deteriorate performance during normal days. Note that this is not the case for ARIMA Day and ARIMA Week that both perform competitively. For both models once the data are seasonally differenced low order polynomials are adequate to fit the time series, hence avoiding the high-dimesional optimisation needed for ETS. Comparing ETS Double with MA Week it can be seen that coding the double seasonal patterns is beneficial for normal periods, but not for outlying ones, where MA Week is the most accurate method. ETS Double outperforms ANN when considering only outlying periods.

Based on these results we can draw the following recommendations: for forecasting normal periods ANN is the most accurate with almost 20% difference from the second best method, providing substantial support for using neural networks in call centre forecasting. However, when evaluating the accuracy on outlying periods the much simpler MA Week is best. In the following section we will attempt to augment ANNs by modelling these outlying periods explicitly. We choose ANNs not only for their good performance, but also for simplicity in extending them to include additional information.

## 4. Modelling call arrival outliers with ANNs

Above we empirically established that ANN is a good candidate for producing accurate call centre arrival forecasts. Based on this we evaluate a number of strategies to code outliers in the inputs of ANNs, which range from merely indicating to the network that an event is taking place to providing information about the shape of the outlier. Consequently, we assume that outliers are known in advance and that all outliers are labelled accordingly. This allows us to compare the alternative strategies proposed by having to encode always the same outlying days and focus solely on assessing their performance. We do not consider additional exogenous or dynamic variables such as days preceding or following a 'special day' and our focus is on modelling the outlying days exclusively. However the contextual informa-

tion of normal or outlying days, i.e. the type and shape of preceding days, is available through the autoregressive inputs of ANNs. In total we consider nine different strategies. All these attempt to capture the functional form of the outlying periods. To further highlight the usefulness of ANNs in modelling such outliers we augment ETS double to include the same inputs and evaluate the relative gains.

### 4.1. Outlier encoding methods

#### 4.1.1. Single Binary Dummy Variable

Using binary dummy variables to capture outliers is standard practice in time series modelling. The binary nature of the dummy contains the information whether an event is occurring or not. In a regression context that simply models an additive shift. However for ANNs the effect is not restricted to additive shifts, or even the same effect when the event is happening, as we discuss in more detail in section 5.1. Therefore, a very simple input for the ANN is to add a single binary dummy whenever an outlier is known to occur, which will be named *ANN Bin1*.

#### 4.1.2. Multiple Binary Dummy Variables

In a linear regression context to model multiple outlying periods that have different periods several binary dummies must be introduced. We can use the same approach with ANNs, where parsimony is traded with aiding the model to fit to the data. To capture a shape that lasts $S$ periods we need $S - 1$ binary dummy variables to eliminate information redundancies in a linear context, resulting in *ANN BinS-1*. Neural networks, being nonlinear, do not need this special treatment resulting in *ANN BinS*. Alternatively, we can introduce a variable selection step to choose how many separate binary dummies need to be retained. Depending if a stepwise or backward approach is used these are named *ANN Bin Step* and *ANN Bin Back* respectively.

#### 4.1.3. Single Integer Dummy Variable

Using a single binary dummy to code functional outliers merely provides the information to the network that an event is taking place and the rest is left to its approximation capabilities. However note that the approximation starts by assuming the effect is identical across the various periods, as the binary dummy is always equal to 1. We can indicate to the network that this is not true by switching the binary dummy with an integer dummy valued: $1, 2, \ldots, S$. Again we rely to the approximation capabilities of the network

16

to capture the exact effect of the outlier. This very parsimonious coding has been shown to be an effective way to capture deterministic seasonality with ANNs (Crone and Kourentzes, 2009). We will refer to this strategy as *ANN Int.*

### 4.1.4. Trigonometric Dummy Variables

Using trigonometric variables to capture patterns with neural networks is common for seasonality. We extend its use to capture patterns of outlying days, using a pair of sine and cosine variables with a daily period. When periods are normal the values of the trigonometric variables are replaced to zero. This way we indicate to the network both the duration of the special events and their timing of occurrence. This strategy requires only two variables and is named *ANN SinCos.*

### 4.1.5. Profile Dummy Variable

Another approach is to explicitly provide the network an indication of the shape of the outlier. A single input is used, the values of which are calculated as the average of all outlying shapes. Again we rely upon the approximation capabilities of the network to fit to the different specific outliers and this strategy is referred to as *ANN Profile.*

### 4.1.6. Model Separately

An alternative would be to model outlying and normal periods independently. In this case we construct two subseries from the original one. All outliers are separated into a new series containing only outliers. Another series is created from the original one, where all outliers are replaced with the seasonal median of the original series normal periods. Replacing outliers with the median is a practical approximation, which is robust to any unknown outliers in the time series. The 'outlier' series is modelled by a separate network. This strategy, which is named *ANN Replace*, requires building two separate networks, but no special new inputs.

### 4.2. Double Seasonal Exponential Smoothing with outlier coding

We augment ETS Double to include the various outlier coding strategies by revising the forecast equation of the model. The seasonality states are calculated as in the normal case, but the measurement equation (6) is modified to include an additional additive term with regressors $\boldsymbol{X}_t\boldsymbol{r}$, where $\boldsymbol{X}_t$

contains all additional variables prescribed by the strategies above at time $t$ and $\boldsymbol{r}$ is a vector containing the respective coefficients for each dummy.

$$y_t = l_{t-1} + s_{t-m_1} + d_{t-m_2} + \phi z_{t-1} + \boldsymbol{X}_t \boldsymbol{r} + e_t. \qquad (15)$$

The estimation of the model is done in the conventional way. For additional details how to include exogenous variables in ETS and evidence of the forecasting performance of the modification, albeit for the single seasonal case, the reader is referred to Hyndman et al. (2008) and Kourentzes and Petropoulos (2015).

### 4.3. Empirical evaluation

Using the same experimental setup as before we compare the performance of the alternative outlier coding strategies. Table 2 provides the average RMAE for overall, normal and outlier periods, both for ANN and ETS Double.

Table 2: Average RMAE of ANN and ETS Double with outlier coding

| Method | ANN | | | ETS Double | | |
|---|---|---|---|---|---|---|
| | Overall | Outlier | Normal | Overall | Outlier | Normal |
| Control | 1.000 | 1.000 | 1.000 | 1.119 | 0.924 | 1.196 |
| Bin1 | 0.976 | 0.931 | 0.992 | 1.132 | 0.965 | 1.193 |
| BinS | 0.903 | **0.844** | 0.927 | 1.610 | 1.190 | 1.790 |
| BinS-1 | **0.898** | 0.850 | **0.919** | 1.604 | 1.190 | 1.782 |
| Bin Step | 0.913 | 0.867 | 0.933 | 1.614 | 1.190 | 1.795 |
| Bin Back | 0.913 | 0.862 | 0.935 | 1.414 | 1.089 | 1.566 |
| Int | 0.946 | 0.926 | 0.954 | 1.124 | 0.944 | 1.192 |
| SinCos | 0.966 | 0.889 | 0.992 | 1.378 | 1.038 | 1.536 |
| Profile | 0.975 | 0.944 | 0.987 | 1.136 | 1.002 | 1.193 |
| Replace | 1.016 | 1.055 | 0.997 | 1.182 | 1.188 | 1.192 |

First let us focus on the ANN results. With the exception of ANN Replace, all other approaches improve upon the base ANN. ANN Replace also improves when evaluated only on normal periods. This demonstrates that ANNs have been able not only to fit the normal periods better, but also learn the behaviour of the outlying periods and predict them relatively accurately in the out-of-sample.

Let us focus our analysis at the performance of the different methods for forecasting normal periods. We anticipate that if the outliers are modelled

accurately, this will have a positive impact in the accuracy of the method for the normal periods, since the parametrisation will not be biased, as discussed above. Comparing the different methods we can observe some distinct behaviours. Methods that use multiple binary dummies (ANN BinS, ANN Bin S-1, ANN Bin Step and ANN Bin Back) perform best, with ANN BinS-1 being the most accurate. The neural network has been able to code the various outliers using in this particular case 47 binary dummies for 48 observations-long days. The flexibility of the network makes it possible to model the complex behaviours of the different types of outliers in a relatively simple way, without needing the modeller to explicitly code different causes or groups of outliers.

ANN Int follows in terms of accuracy. It was assumed that the network would benefit by providing the information that the periods of the outlier have different values. Comparing ANN Int with ANN Bin there is support for this.

ANN SinCos has weak performance. To explain this we need to consider the nature of the inputs. When an outlier was observed a complete period of a sine and a cosine are provided as inputs. These variables are equal to zero for all other periods. A zero input is expected to have no effect in a linear regression, but as discussed above a neural network can code multiple responses with the same input.

Finally ANN Replace improves upon the base ANN, but only marginally. This shows that the seasonal median used to replace the outlying values in the'normal subseries' was not adequate. Although one can argue about different ways to replace outlying values, letting the neural network simultaneously estimate normal and outlying periods is simpler and performs well as the rest of the results suggest.

The results for the the outlying periods are similar. Note that ANN Replace performs worse than the base ANN that has no special treatment for outliers. To forecast the outlying periods ANN Replace fits a network on a series comprised solely by outliers. The empirical evidence suggests that the lack of systematic patterns does not allow the networks to train adequately.

ANN BinS and ANN BinS-1 are simpler than ANN Bin Step and ANN Bin Back, and therefore should be preferred since their is no observed loss of accuracy. Between the two there are marginal differences and therefore we recommend using the most parsimonious so as to make the computational cost of the network smaller. Consider that even a single additional input requires $H$ additional weights to be trained, the number of the hidden nodes.

The rest of table 2 provides the results for ETS Double. In this case we do not see the performance gains observed by ANN. This can be explained by the way that the additional information is handled by ETS Double. The model is capable of capturing unique additive linear effects for each additional regressor, as quantified by $r$. This lacks the flexibility of the nonlinear ANN and ETS Double would require much more complex handling of outliers. We argue that the resulting modelling simplicity of outliers with ANNs counterbalances the additional complexity of the method itself over standard statistical models.

Concluding, table 3 provides the percentage gains in accuracy of the recommended ANN BinS-1 over the best performing statistical benchmarks from table 1. We can observe that overall accuracy gains are about 20%, with on average more than 5% gains in accuracy during periods of outliers.

Table 3: Percent gains in accuracy of ANN BinS-1

| Method | Overall | Outlier | Normal |
|---|---|---|---|
| MA Week | +22.7% | +4.7% | +27.1% |
| ETS Double | +18.4% | +6.2% | +22.0% |

A more detailed comparison of the top performing members of each model family is provided in Fig. 4. We observe that ANN with no outlier coding are consistently more accurate than other univariate alternatives across all forecast horizons. The inclusion of such information (ANN BinS-1) in turn consistently outperforms the univariate ANN. The differences for short horizons between the forecasts are small, with the exception of Naïve Week and MA Week the RMAE of which is poor for short horizons but improves as the errors of the other forecasts increases for longer horizons.

## 5. Discussion

In this section we consider why ANNs work and the implications for theory and practice. We argue here that their flexibility makes them particularly suited for modelling this type of outliers and data. Considering Eq. (10) one can observe that the various network inputs can interact. This permits ANN to be 'context aware' when capturing outliers. This reduces the need for the call centre analyst and forecaster to specify complex and often unknown behaviour of outlying and 'special' days.
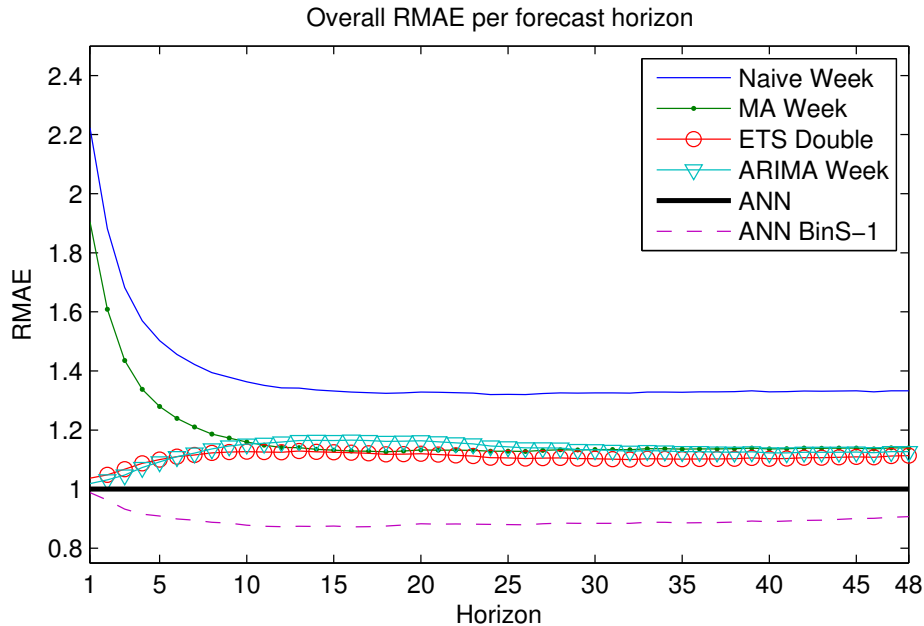
Figure 4: RMAE per horizon for top performing forecasts.

## 5.1. Why ANNs work: An illustrative example

To better illustrate this property of ANNs consider the example of a seasonal time series with two outliers, one approximately during the season minimum and one during the maximum. Both outliers change the value towards the mean of the time series, therefore require adjustment of the baseline towards the opposite direction. We construct ANNs with two neurons in the hidden layer, following the setup described earlier. We fit one network with appropriate autoregressive inputs and a single binary dummy to indicate when the outlier is taking place and a network without the additional dummy variable. The resulting fits can be seen in Fig. 5. Observe that the ANN without the additional information has poor fit during the outlying periods, while the opposite is true for the ANN with the dummy variable. Note that the network is capable of fitting well to the two different outliers, that require opposite adjustments, using only a single dummy variable.

Due to the small size of the network we can fully observe how the output is calculated during the two outlying periods. Fig. 6 visualises the two neurons in the hidden layer and the output node. The input weights and the respective inputs and outputs are provided. The network has three inputs,
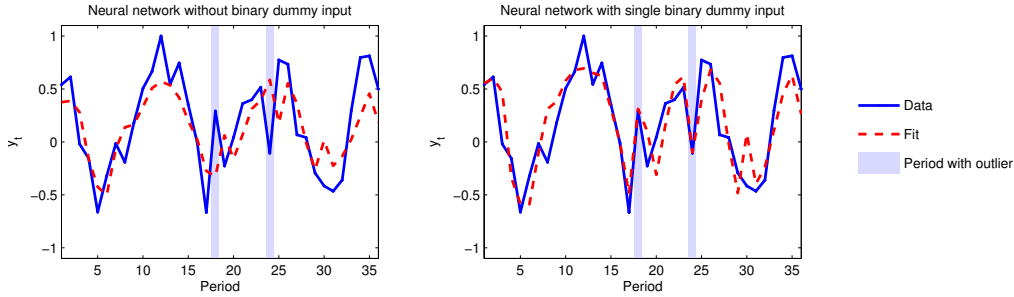
21

Figure 5: Resulting fit of ANN without and with a single binary dummy to indicate the presence of outliers.

two autoregressive lags of order 1 and 12 and a single binary dummy. To highlight the impact of the dummy the inputs and outputs with and without the dummy for the two outlying periods are provided. Note that during the remaining normal periods the value of the binary dummy is equal to zero and therefore does not affect the construction of the forecasts.
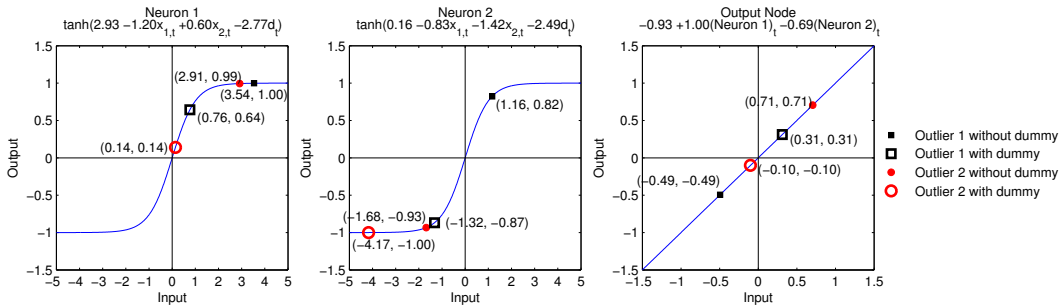


Figure 6: ANN neurons inputs and outputs for the outlying periods.

Focusing on the first neuron we observe that without the dummy variable the input of the hyperbolic tangent transfer function is equal to the sum of the autoregressive inputs multiplied by their respective weights, plus a constant term. For both outlying periods the resulting value is relatively high and at the highly nonlinear region of the hyperbolic tangent. On the other hand when the binary dummy is considered, the inputs are shifted towards the linear part of the neuron. Observe that the shift for the first outlier (square) is $-0.36$ while for the second (circle) is $-0.86$, due to the nonlinearity of the transfer function.

22

For the second neuron the inputs for the second outlier remain at the nonlinear region with only small changes in the output, as the shift of $-2.49$ cause by the binary dummy has minimal effect due to the shape of the hyperbolic tangent transfer function. Contrary, for the first outlier the binary dummy shifts the inputs to a different nonlinear behaviour, resulting to a substantial change in the output. The total change of the output for the first outlier is just $-0.07$, while for the second is $-1.09$. The relatively massive difference in the outputted value is due to the sum of the weighted autoregressive inputs before the effect of the binary dummy is considered, pushing the input to a different region of the nonlinearity and thus causing the dummy to have a different effect.

The linear output node combines the outputs from the two neurons in the hidden node and shifts them by a constant value. Observe that the weight of the first neuron is approximately equal to 1. The second neuron has a weight of $-0.69$, so as both outputs of the second neuron were negative they will contribute with a positive shift to the final output. For the first outlier (square) the total shift caused by the inclusion of the dummy variable results in an upward shift. This is appropriate as the first outlier is higher than the anticipated value of the series. On the other hand the total shift for the second outlier is negative, effectively reducing the output value, resulting in an opposite shift as it is appropriate for the time series in hand. This behaviour arises from the nonlinear behaviour of the neurons and the interaction of the inputs of the ANN, something that is impossible with other conventional statistical models in section 3, as demonstrated in table 2 as well.

This illustrates that ANNs have a substantial advantage in being able to model economically, in terms of number of inputs, complex outlying behaviours. Introducing only a limited number of dummy variables is enough to capture the multiple different shapes and magnitudes of outlying days avoiding the need to resort to overcomplicated outlier encoding or treatment methods. This is very useful for modelling high frequency outliers that may last several periods as is the case for the call centre time series. In such cases the ANN is aware of any interactions with the baseline values, as well as any proceeding outlying values.

### 5.2. Implications for practice

The problem of modelling outliers and special days for high frequency data remains a major challenge in call centre forecasting and OR practice in general. In the literature there is very limited research on the topic. Due

23

to the nature of the data conventional outlier modelling methods are not applicable, which has led practitioners to develop various ad-hoc approaches in lack of best practices. Our work is attempting to contribute to this by evaluating several strategies to model such days.

At the same time we investigate the performance of various forecasting methods for predicting the volume of call arrivals. Although this is not the first study to do this, it differs in two aspects relevant for practice: (i) it extends the types of methods considered to include ANN, a flexible nonlinear forecasting method; (ii) evaluates each method by considering a modelling methodology that can be fully automated.

The first aspect turns out to be significant, as we find ANNs to be uniquely capable in modelling and predicting outlying days with only simple extensions of the input vector. This limits the need for ad-hoc unproven approaches that practitioners often have to resort to in dealing with such days. We argue that this simplifies the forecasting task.

The second aspect is very relevant for practice as automation is important for organisations. Forecasts have to be scalable, adaptive to the problem and dataset at hand and relatively insensitive to different levels of the user's statistical expertise. Organisations typically give more importance to the domain knowledge of the expert, rather than statistical skills. The various alternative methods that were considered in this analysis can all be implemented in a fully automatic way and all are arguably feasible implementations for companies. Although statistical expertise cannot be replaced, solutions that do not require an extensive setup or customisation from users can be more widely and easily used in practice.

We find that ANNs are not only producing accurate call arrivals volume predictions, but also capture outliers effectively. Therefore a single forecasting method is capable of handling both normal and special days. We anticipate that this can eventually lead in a simplification of the forecasting process itself. Although as a method ANNs are surely more complex than other alternatives evaluated here, such as Seasonal Moving Averages that have been found to perform well, once seen as a holistic forecasting solution the complexity of ANNs is offset by the simplification of the forecasting process. Last but not least, the expert no longer needs to manually treat and forecast special days, thus having more time to simply mark such days for the statistical forecast and focus on adjusting any predictions when additional useful information is available. Therefore we argue that a simplified forecasting process allows experts to add more value through their domain

knowledge skills. Extending this argument, it raises an interesting question for the design and software implementation of forecasting support systems. We already argued that methodologies that can be automated are crucial for practice. Their benefits will only be fully materialised if the forecasting support systems provide the appropriate support to the expert user in terms of information provided and feedback, as well as inputs to the system.

Nonetheless, ANNs have been criticised for being 'black-box'. Our recommended method does not provide descriptive insights to experts. It primarily offers predictive improvements and it is still left up to the expert to highlight the behaviour of a particular special event. To this end ANNs can support experts by means of sensitivity analysis, where one can simulate the effect on the forecasts of various events and special days. This is far from a complete solution, but it is offering a new direction for further research and a starting point for practitioners to model such days. It is natural that the modelling objective will indicate the most appropriate method for practitioners.

Finally, a relevant discussion point is the generality of the strategies to code outliers evaluated here to other high frequency time series forecasting problems. Although one would need to empirically verify their effectiveness, we anticipate that applications with similar data characteristics, such as television and online viewership, electricity load, server load of online activities and so on, would benefit, thus having potential for wider practical usefulness.


## 6. Conclusions and future research

In this paper we have gone beyond previous studies on forecasting call centre arrival data to consider the effect of outlying periods and how to best forecast call arrivals when these are present. We find ANNs to be the most accurate of all methods in our empirical evaluation. However, when looking only during outlying periods the simpler double seasonal exponential smoothing and the seasonal weekly moving average perform better.

Motivated by the flexibility of ANNs we go on to show that their performance can be substantially improved by explicitly modelling outlying periods, with improvements in accuracy of about 20% overall, and more than 5% gains in accuracy during outlying periods over the best performing alternatives. The use of multiple binary dummies is found to yield the best performance of all methods for the chosen dataset, while ANN Bin S-1 is chosen over ANN BinS due to having fewer free parameters and marginally superior accuracy. We argue that even though ANNs are relatively more

complex than conventional statistical methods, the simplicity of modelling outlying call arrivals periods makes them attractive.

However more work is needed in this area. Specifically it may be the case that not all outlying periods are labelled correctly. Future work should aim to introduce methods for automatically identifying and validating the presence outlying patterns, especially given the high frequency nature of call centre data, as well as the number of different call types that maybe monitored.

Furthermore, in this study we only focused on evaluating point forecasts of call arrival volumes. Naturally it is possible to make use of empirical approaches to calculate prediction error distributions for the methods in this paper, when analytical ones are not available. Density forecasts of arrival rates are useful for supporting staffing decisions and accounting for uncertainty in call arrival volumes (Gans et al., 2003; Taylor, 2008a). Taylor (2012) show how density forecasts can be produced using models based on exponential smoothing together with assumptions of Poisson arrivals, while Weinberg et al. (2007) generate density forecasts using a multiplicative Gaussian model estimated within a Bayesian framework. We see the incorporation of our work on outlier modelling into density forecast models of call centre arrival rates as useful future research.

### References

Adya, M., Collopy, F., 1998. How effective are neural networks at forecasting and prediction? a review and evaluation. J. Forecasting 17, 481–495.

Aksin, Z., Armony, M., Mehrotra, V., 2007. The modern call center: A multi-disciplinary perspective on operations management research. Production and Operations Management 16 (6), 665–688.

Aldor-Noiman, S., Feigin, P. D., Mandelbaum, A., 2009. Workload forecasting for a call center: Methodology and a case study. The Annals of Applied Statistics, 1403–1447.

Andrews, B. H., Cunningham, S. M., 1995. Ll bean improves call-center forecasting. Interfaces 25 (6), 1–13.

Antipov, A., Meade, N., 2002. Forecasting call frequency at a financial services call centre. Journal of the Operational Research Society 53 (9), 953–960.

Armstrong, J. S., 2006. Findings from evidence-based forecasting: Methods for reducing forecast error. International Journal of Forecasting 22 (3), 583–598.

Avramidis, A. N., Deslauriers, A., L'Ecuyer, P., 2004. Modeling daily arrivals to a telephone call center. Management Science 50 (7), 896–908.

Balaguer, E., Palomares, A., Soria, E., Martín-Guerrero, J. D., 2008. Predicting service request in support centers based on nonlinear dynamics, arma modeling and neural networks. Expert Systems with Applications 34 (1), 665–672.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., 2008. Time Series Analysis: Forecasting and Control, 4th Edition. Wiley, New Jersey.

Brown, L., Gans, N., Mandelbaum, A., Sakov, A., Shen, H., Zeltyn, S., Zhao, L., 2005. Statistical analysis of a telephone call center: A queueing-science perspective. Journal of the American statistical association 100 (469), 36–50.

Chatfield, C., 2013. The analysis of time series: an introduction. CRC press.

Conejo, A. J., Contreras, J., Espínola, R., Plazas, M. A., 2005. Forecasting electricity prices for a day-ahead pool-based electric energy market. International Journal of Forecasting 21 (3), 435–462.

Crone, S. F., Hibon, M., Nikolopoulos, K., 2011. Advances in forecasting with neural networks? empirical evidence from the nn3 competition on time series prediction. International Journal of Forecasting 27 (3), 635–660.

Crone, S. F., Kourentzes, N., 2009. Forecasting seasonal time series with multilayer perceptrons-an empirical evaluation of input vector specifications for deterministic seasonality. In: DMIN. pp. 232–238.

Crone, S. F., Kourentzes, N., 2010. Feature selection for time series prediction–a combined filter and wrapper approach for neural networks. Neurocomputing 73 (10), 1923–1936.

Davydenko, A., Fildes, R., 2013. Measuring forecasting accuracy: The case of judgmental adjustments to SKU-level demand forecasts. International Journal of Forecasting 29 (3), 510–522.

De Livera, A. M., Hyndman, R. J., Snyder, R. D., 2011. Forecasting time series with complex seasonal patterns using exponential smoothing. Journal of the American Statistical Association 106 (496), 1513–1527.

Draper, N. R., John, J. A., 1981. Influential observations and outliers in regression. Technometrics 23 (1), 21–26.

Gans, N., Koole, G., Mandelbaum, A., 2003. Telephone call centers: Tutorial, review, and research prospects. Manufacturing & Service Operations Management 5 (2), 79–141.

Gardner, E. S., 2006. Exponential smoothing: The state of the art - part II. International Journal of Forecasting 22 (4), 637 – 666.

Gelper, S., Fried, R., Croux, C., 2010. Robust forecasting with exponential and holt-winters smoothing. Journal of forecasting 29 (3), 285–300.

Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., Vahid-Araghi, F., 2008. Forecasting time series with multiple seasonal patterns. European Journal of Operational Research 191 (1), 207–222.

Hamid, S. A., Iqbal, Z., 2004. Using neural networks for forecasting volatility of s&p 500 index futures prices. Journal of Business Research 57 (10), 1116–1125.

Hippert, H. S., Bunn, D. W., Souza, R. C., 2005. Large neural networks for electricity load forecasting: Are they overfitted? International Journal of forecasting 21 (3), 425–434.

Hornik, K., 1991. Approximation capabilities of multilayer feedforward networks. Neural networks 4 (2), 251–257.

Hyndman, R. J., Khandakar, Y., 7 2008. Automatic time series forecasting: The forecast package for R. Journal of Statistical Software 27 (3), 1–22.

Hyndman, R. J., Koehler, A. B., Ord, J. K., Snyder, R. D., 2008. Forecasting with Exponential Smoothing: The State Space Approach. Springer Verlag, Berlin.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., Grose, S., 2002. A state space framework for automatic forecasting using exponential smoothing methods. International Journal of Forecasting 18 (3), 439–454.

Hyndman, R. J., Shang, H. L., 2009. Forecasting functional time series. Journal of the Korean Statistical Society 38 (3), 199–211.

Hyndman, R. J., Shang, H. L., 2010. Rainbow plots, bagplots, and boxplots for functional data. Journal of Computational and Graphical Statistics 19 (1).

Ibrahim, R., L'Ecuyer, P., 2013. Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models. Manufacturing & Service Operations Management 15 (1), 72–85.

Jongbloed, G., Koole, G., 2001. Managing uncertainty in call centres using poisson mixtures. Applied Stochastic Models in Business and Industry 17 (4), 307–318.

Khotanzad, A., Afkhami-Rohani, R., Maratukulam, D., 1998. Annstlf-artificial neural network short-term load forecaster generation three. Power Systems, IEEE Transactions on 13 (4), 1413–1422.

Koehler, A. B., Snyder, R. D., Ord, J. K., Beaumont, A., 2012. A study of outliers in the exponential smoothing approach to forecasting. International Journal of Forecasting 28 (2), 477–484.

Kolassa, S., 2011. Combining exponential smoothing forecasts using Akaike weights. International Journal of Forecasting 27 (2), 238–251.

Kourentzes, N., 2011. Modelling functional outliers for high frequency time series forecasting with neural networks. In: Stahlbock, R. (Ed.), The 2011 International Conference on Data Mining, DMIN 2011. CSREA Press, pp. 313–319.

Kourentzes, N., Barrow, D. K., Crone, S. F., 2014a. Neural network ensemble operators for time series forecasting. Expert Systems with Applications 41 (9), 4235–4244.

Kourentzes, N., Crone, S. F., 2010. Frequency independent automatic input variable selection for neural networks for forecasting. In: Neural Networks (IJCNN), The 2010 International Joint Conference on. IEEE, pp. 1–8.

Kourentzes, N., Petropoulos, F., 2015. Forecasting with multivariate temporal aggregation: The case of promotional modelling. International Journal of Production Economics.

Kourentzes, N., Petropoulos, F., Trapero, J. R., 2014b. Improving forecasting by estimating time series structural components across multiple frequencies. International Journal of Forecasting 30 (2), 291–302.

Lee Willis, H., Northcote-Green, J. E. D., 1983. Spatial electric load forecasting: a tutorial review. Proceedings of the IEEE 71 (2), 232–253.

Li, M., Huang, L., Gong, L., 2011. Research on the challenges and solutions of design large-scale call center intelligent scheduling system. Procedia Engineering 15, 2359–2363.

Millan-Ruiz, D., Hidalgo, J. I., 2010. A memetic algorithm for workforce distribution in dynamic multi-skill call centres. In: Evolutionary Computation in Combinatorial Optimization. Springer, pp. 178–189.

Møller, M. F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural networks 6 (4), 525–533.

Ord, J. K., Fildes, R., 2013. Principles of Business Forecasting. South-Western Cengage Learning, Mason, Ohio.

Pacheco, J., Millan-Ruiz, D., Velez, J. L., 2009. Neural networks for forecasting in a multi-skill call centre. In: Engineering Applications of Neural Networks. Springer, pp. 291–300.

Petropoulos, F., Makridakis, S., Assimakopoulos, V., Nikolopoulos, K., 2014. horses for courses in demand forecasting. European Journal of Operational Research 237 (1), 152–163.

Ramsay, J. O., Silverman, B. W., 2002. Applied functional data analysis: methods and case studies. Vol. 77. Springer New York.

Shen, H., 2009. On modeling and forecasting time series of smooth curves. Technometrics 51 (3).

Shen, H., Huang, J. Z., 2005. Analysis of call centre arrival data using singular value decomposition. Applied Stochastic Models in Business and Industry 21 (3), 251–263.

Silverman, B. W., Ramsay, J. O., 2005. Functional Data Analysis. Springer.

Tandberg, D., Easom, L. J., Qualls, C., 1995. Time series forecasts of poison center call volume. Clinical Toxicology 33 (1), 11–18.

Taylor, J. W., 2003. Short-term electricity demand forecasting using double seasonal exponential smoothing. Journal of the Operational Research Society 54 (8), 799–805.

Taylor, J. W., 2008a. A comparison of univariate time series methods for forecasting intraday arrivals at a call center. Management Science 54 (2), 253–265.

Taylor, J. W., 2008b. Exponentially weighted information criteria for selecting among forecasting models. International Journal of Forecasting 24 (3), 513–524.

Taylor, J. W., 2010a. Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles. International Journal of Forecasting 26 (4), 627–646.

Taylor, J. W., 2010b. Triple seasonal methods for short-term electricity demand forecasting. European Journal of Operational Research 204 (1), 139–152.

Taylor, J. W., 2012. Density forecasting of intraday call center arrivals using models based on exponential smoothing. Management Science 58 (3), 534–549.

Taylor, J. W., De Menezes, L. M., McSharry, P. E., 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. International Journal of Forecasting 22 (1), 1–16.

Taylor, J. W., Snyder, R. D., 2012. Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. Omega 40 (6), 748–757.

Temraz, H. L., Salama, M. M. A., Chikhani, A. Y., 1997. Review of electric load forecasting methods. In: Electrical and Computer Engineering, 1997. Engineering Innovation: Voyage of Discovery. IEEE 1997 Canadian Conference on. Vol. 1. IEEE, pp. 289–292.

Tsay, R. S., 1989. Testing and modeling threshold autoregressive processes. Journal of the American Statistical Association 84 (405), 231–240.

Weinberg, J., Brown, L. D., Stroud, J. R., 2007. Bayesian forecasting of an inhomogeneous poisson process with applications to call center data. Journal of the American Statistical Association 102 (480), 1185–1198.

Zhang, G., Patuwo, B. E., Hu, M. Y., 1998. Forecasting with artificial neural networks:: The state of the art. International journal of forecasting 14 (1), 35–62.