

Rausch, Theresa Maria; Albrecht, Tobias; Baier, Daniel

Article — Published Version

Beyond the beaten paths of forecasting call center arrivals: on the use of dynamic harmonic regression with predictor variables

Journal of Business Economics

Provided in Cooperation with:

Springer Nature

Suggested Citation: Rausch, Theresa Maria; Albrecht, Tobias; Baier, Daniel (2021) : Beyond the beaten paths of forecasting call center arrivals: on the use of dynamic harmonic regression with predictor variables, Journal of Business Economics, ISSN 1861-8928, Springer, Berlin, Heidelberg, Vol. 92, Iss. 4, pp. 675-706,
<https://doi.org/10.1007/s11573-021-01075-4>

This Version is available at:

<https://hdl.handle.net/10419/286906>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Beyond the beaten paths of forecasting call center arrivals: on the use of dynamic harmonic regression with predictor variables

Theresa Maria Rausch¹ · Tobias Albrecht² · Daniel Baier¹

Accepted: 1 December 2021 / Published online: 17 December 2021
© The Author(s) 2021

Abstract

Modern call centers require precise forecasts of call and e-mail arrivals to optimize staffing decisions and to ensure high customer satisfaction through short waiting times and the availability of qualified agents. In the dynamic environment of multi-channel customer contact, organizational decision-makers often rely on robust but simplistic forecasting methods. Although forecasting literature indicates that incorporating additional information into time series predictions adds value by improving model performance, extant research in the call center domain barely considers the potential of sophisticated multivariate models. Hence, with an extended dynamic harmonic regression (DHR) approach, this study proposes a new reliable method for call center arrivals' forecasting that is able to capture the dynamics of a time series and to include contextual information in form of predictor variables. The study evaluates the predictive potential of the approach on the call and e-mail arrival series of a leading German online retailer comprising 174 weeks of data. The analysis involves time series cross-validation with an expanding rolling window over 52 weeks and comprises established time series as well as machine learning models as benchmarks. The multivariate DHR model outperforms the compared models with regard to forecast accuracy for a broad spectrum of lead times. This study further gives contextual insights into the selection and optimal implementation of marketing-relevant predictor variables such as catalog releases, mail as well as postal reminders, or billing cycles.

Keywords Forecasting · Call center arrivals · Dynamic harmonic regression · Time series analysis · Machine learning · Customer relationship management

JEL Classification C22 · C53 · L81 · M30

✉ Theresa Maria Rausch
theresa.rausch@uni-bayreuth.de

Extended author information available on the last page of the article

1 Introduction

In the retail industry, typical stages along the customer journey like order taking, after-sales service, and complaint resolution can easily be completed online (Gensler et al. 2012; Verhoef et al. 2015). Nevertheless, approximately 70% of customers still prefer to interact with a human counterpart for customer service requests (Sitel Group 2018). For many organizations, call centers constitute the main or only point of human-to-human interaction with their customers. Thus, call centers are an essential communication channel for businesses and an important customer touchpoint across many industries (Aksin et al. 2007).

High perceived service quality at this customer interface contributes greatly to customer loyalty and is determined by short waiting times as well as the call experience and service outcome itself (Brady and Cronin 2001; Parasuraman et al. 1985; Zeithaml et al. 1996). Hence, to provide the correct number of call center agents as customer service representatives at the right time and to evaluate their required areas of expertise, call arrival volumes in different queues have to be predicted¹ reliably in advance. In this regard, preceding literature in the fields of operations management and forecasting so far focused on optimizing the opposite tendency of staffing costs and customer waiting times by enhancing forecast accuracy of predominant time series models (Dean 2007; Gans et al. 2003). Methods like autoregressive integrated moving average (ARIMA) (Andrews and Cunningham 1995; Barrow 2016; Mabert 1985; Thompson and Tiao 1971), exponential smoothing (Taylor 2003, 2008, 2012), or dynamic harmonic regression (DHR) (Young et al. 1999; Young 1999) have traditionally been established as standard-setting approaches. However, such time series models generate predictions based on the time series' previous values without including any contextual data or other additional information available.

Meanwhile, new methods apart from time series models were developed and investigated to predict call center arrivals with high accuracy (Albrecht et al. 2021; Rausch and Albrecht 2020). Regression models (Aldor-Noiman et al. 2009; Brown et al. 2005; Ibrahim et al. 2016; Ibrahim and L'Ecuyer 2013) and machine learning (ML) approaches like random forest (RF) and neural networks (Barrow 2016; Jalal et al. 2016) were found to yield accurate predictions by including meaningful predictor variables. Incorporating contextual data into call arrival forecasts not only positively affects prediction accuracy but also allows for a more customer-centric perspective in call center forecasting. By including information on customer motives and behavior, valuable marketing insights are gained. Thus, prior research recommended to model predictor variables (Taylor 2008).

Hence, extant forecasting literature still disagrees on the best-performing model type for call arrival forecasting. Simultaneously, from a conceptual point of view, previous research did not explicitly investigate whether incorporating the benefits of both time series approaches as well as regression and ML models into one model yields an unrecognized predictive potential. We therefore contribute to literature

¹ The terms predicting and forecasting will be utilized interchangeably in this paper.

by proposing a new method for call center arrival forecasting, which combines the strengths of those model types—on the one hand by capturing the dynamics of the time series and on the other hand by including predictor variables. We hypothesize that this approach will lead to an increase in predictive performance and, simultaneously, will entail advantages for practical use. We extend the established Dynamic Harmonic Regression (DHR) model, which utilizes a sum of sinusoidal terms (i.e., Fourier terms) as predictors to handle periodic seasonality and an ARIMA error to capture short-term dynamics, by including predictor variables in the considered information space to generate predictions. We test the predictive potential of our approach with two different call and e-mail arrival series of a leading German online retailer comprising 174 weeks of data. We compare our proposed model to different established time series and ML approaches with time series cross-validation and an expanding rolling window. We further extend knowledge on suitable predictor variables in a marketing context, which has been neglected by prior research, as most datasets do not include such contextual information.

The remainder of this paper is structured as follows: Sect. 2 reviews related work on call center arrival forecasting approaches. In Sect. 3, we derive our DHR model including predictor variables. Sections 4 and 5 analyze the customer support call and e-mail arrival series respectively by presenting the preliminaries of both datasets, the experimental design, and the analysis results. Section 6 discusses the results with regard to the practical implications as well as the theoretical contribution of the study before it reflects the limitations and provides guidance for future research. The paper concludes with a concise summary of its principal points in Sect. 7.

2 Related work

Call center call arrivals are count data and hence, discrete data restricted to non-negative integers. Therefore, a common model utilized for call arrivals' forecasts is a Poisson arrival process (see e.g., Aksin et al. 2007; Cezik and L'Ecuyer 2008; Gans et al. 2003; Taylor 2012). However, one key feature of call center arrivals, which is not aligning with the preceding Poisson assumption, is their time dependence: call arrival counts exhibit obvious patterns (or seasonalities respectively) which are repeating itself sub-daily (e.g., half-hourly, hourly), daily, weekly, or yearly (Brown et al. 2005; Ibrahim et al. 2016). Thus, literature frequently draws on time series analysis to forecast call center arrivals assuming them to be a sequence of dependent, contiguous observations which are made continuously over a certain time interval (Brockwell and Davis 2002).

Box and Jenkins (1970) developed a non-seasonal ARIMA (p, d, q) which assumes that—with d degrees differenced—the time series y'_t at time t is dependent on past values p periods apart (autoregressive part) and is related to a finite number q of preceding forecast errors ε (moving average part). The ARIMA model—or reduced models containing only sub-components respectively—are among the most common approaches to predict future call arrivals. The fields of application include e.g. a public telephone company (Thompson and Tiao 1971), an emergency line (Mabert 1985), banks in the US, UK, and Israel (Barrow

2016), or a retailer for outdoor goods and apparel (Andrews and Cunningham 1995). In the latter case, additional contextual covariates (i.e., catalog mailings and holidays) were modeled to enhance forecast accuracy (Andrews and Cunningham 1995). Bianchi et al. (1998) forecasted call center arrivals for telemarketing centers and found ARIMA modeling to be superior to Holt-Winters' exponentially weighted moving average.

Exponential smoothing methods are predicting values by weighting averages of previous observations with exponentially decaying weights the further the observations lie in the past (Holt 2004; Winters 1960). Extensions of the Holt-Winters' approach comprised double seasonality (Taylor 2003), a Poisson count data model with gamma-distributed stochastic arrival rate (Taylor 2012), and robust exponential smoothing (Taylor 2008). For relatively short forecast horizons (up to 6 days), Taylor (2008) found the Holt-Winters' extension to outperform established approaches such as seasonal autoregressive moving average (ARMA). A novel subclass of exponential smoothing models are innovation state space models that add an error term to exponential smoothing models yielding the label Error, Trend, Seasonal (ETS) (\cdot, \cdot, \cdot) for each state space model (Hyndman et al. 2002). The single components can be defined as Error = {Additive (A), Multiplicative (M)}, Trend = {None (N), A} and Seasonal = {N, A, M}. ETS models were found to be both superior (Hyndman et al. 2002) as well as inferior (Barrow and Kourentzes 2018) to ARIMA in a call center forecasting context. De Livera et al. (2011) extended the basic ETS model which allows the seasonalities to slowly change over time by including Fourier terms for a trigonometric representation of seasonality and a Box-Cox transformation. The model's key features trigonometric seasonality, Box-Cox transformation, ARMA errors, trend and seasonal components yield the acronym TBATS.

The random walk (RW) method is an easy-to-implement time series approach frequently used as a benchmark model. Essentially, its forecasts equal the last actual value or observation respectively. By including the drift parameter c , the model additionally captures the average of changes between consecutive observations. Despite its simplicity, its performance is to some extent comparable—but not superior—to established methods in many experimental settings (see e.g., Taylor 2008).

Aside from time dependence, another key property of call center arrivals is their overdispersion, i.e., the variance of an arrival count per time period usually exceeds its mean (Aldor-Noiman et al. 2009; Avramidis et al. 2004; Jongbloed and Koole 2001), which is not consistent with the Poisson assumption. One option to deal with overdispersion is to assume the Poisson arrival process as doubly stochastic with random arrival rate (Whitt 1999). Literature then drew on the root-unroot variance stabilizing data transformation for Poisson data assuming that with a large number of calls per interval, the square-root transformed counts are roughly normally distributed (Aldor-Noiman et al. 2009; Brown et al. 2005, 2010). Normality is then exploited to fit Gaussian linear models. Particularly, linear fixed-effects (Ibrahim and L'Ecuyer 2013; Shen and Huang 2008; Weinberg et al. 2007), random-effects, and mixed-effects models (Aldor-Noiman et al. 2009; Brown et al. 2005; Ibrahim et al. 2016; Ibrahim and L'Ecuyer 2013) were utilized subsequently. Mostly, fixed effects comprise the effect of the day of the week and the time of the day and their interaction, whereas random effects capture the daily volume deviation of the fixed

weekday effect. In call center forecasting, Aldor-Noiman et al. (2009) further modeled billing cycles and delivery periods as additional covariates.

Recently, research started to investigate beyond time series and regression models. In the context of forecasting, approaches within the ML paradigm are characterized by detaching from distributional assumptions while, at the same time, promising immense predictive performance in a variety of application areas (Carbonneau et al. 2008). Rausch and Albrecht (2020) included RF as a powerful ML method in their comparison of novel time series and regression models for call center arrivals forecasting. RF was found to yield higher prediction accuracy for nearly all of the considered lead time constellations. Similar results were gathered in an extensive ML comparison study by Albrecht et al. (2021). Besides, artificial neural networks such as multilayer perceptrons (Barrow 2016) and recurrent neural networks (Jalal et al. 2016) attracted increasing attention. Barrow and Kourentzes (2018) found artificial neural networks to outperform traditional models for model complex outliers in call center arrival forecasting. However, call center forecasting research using ML approaches is still in its infancy.

While time series models generate predictions based on previous values in the time series but generally do not capture additional information, aforementioned ML and regression models allow for the inclusion of predictor variables. As these parameter values are typically available for both the past (i.e., the training data) and the future (i.e., the predictions and the test data), ex-post forecasts can be created. Such data specifying a forecasting period in the future at the time of the prediction may include, for instance, date-related information or data on scheduled customer contact activities. In contrast, time series models' forecasts are solely based on information available at the time of the prediction, i.e., the time series' historical values are used to generate ex-ante forecasts (Taylor 2008). This difference regarding the forecasting method was previously found to significantly affect a models' ability to maintain stable prediction accuracy with varying lead time as the inclusion of predictor variables is assumed to make forecasting call center arrivals more robust and accurate (Rausch and Albrecht 2020). On the other hand, ex-post forecasting models' lacking ability of capturing information of a time series' course and dynamics is supposed to prevent a significant further increase in model performance (Barrow 2016).

In case of ex-post forecasting models, including meaningful context factors in the form of predictor variables is critical as their informative value strongly affects prediction accuracy (Andrews and Cunningham 1995). Previous literature identified an extensive list of possible variables that have been observed to affect arrival volumes. Data specifying date-related patterns such as variables indicating the time of day, the day of the week and holidays are widely used (Ibrahim et al. 2016; Ibrahim and L'Ecuyer 2013; Shen and Huang 2008; Weinberg et al. 2007). Additionally, information regarding customer contact activities on the part of the company like variables revealing upcoming billing cycles, delivery periods and catalog mailings has been investigated (Aldor-Noiman et al. 2009). However, their effect on call center arrival volumes has only been examined for a fixed point of prediction so that findings on the influence on customer call behavior are vague and do not allow for a thorough understanding of the relation. In this regard, capturing the temporal effect of influential factors over time to enable the estimation of short-time and medium-term effects

as well as to assess their interrelation with data-related factors such as holidays and weekends is needed. This would lead to a better transferability of research into the complex and dynamic environment of practical call center arrival forecasting and, at the same time, provide valuable insights into the effect of customer contact activities on customer behavior.

3 Dynamic harmonic regression with predictor variables

The methods used in previous call center forecasting studies presented in Sect. 2 only allow for the inclusion of information from past observations of a time series or the incorporation of external data from predictor variables. Hence, either contextual information possibly stemming from predictor variables or information extracted from time series dynamics is lost. The latter applies to ordinary regression models of form

$$y_t = \beta_0 + \beta_1 x_t + \cdots + \beta_i x_{i,t} + N_t \quad (1)$$

with i predictor variables $x_{i,t}$ at time t and the time series' value y_t at time t . The error term is mostly assumed to be a set of zero-mean and normally distributed white noise random shocks a_t (Pankratz 1991). Thus, if N_t in (1) has mean zero and is normally distributed white noise, then $N_t = a_t$. However, it is problematic estimating ordinary regression models of (1) with time series data (Pankratz 1991). As stated earlier, regression models cannot capture previous dynamics of a time series: E.g., the error term might be autocorrelated, i.e., N_t is related to its previous values N_{t-1} , N_{t-2} , etc., i.e.,

$$N_t = \phi_1 N_{t-1} + a_t \quad (2)$$

with coefficient ϕ_1 and random shock component a_t . Alternatively,

$$N_t = a_t - \theta_1 a_{t-1} \quad (3)$$

with coefficient θ_1 and a_{t-1} being the random shock component of N_{t-1} . Thereby, Eq. (2) represents an autoregressive process whereas Eq. (3) displays a moving average process. Hence, combining both equations yields an ARIMA process. If we allow the error term N_t in (1) to contain autocorrelation of (2) and (3), we obtain a dynamic regression model

$$y_t = \beta_0 + \beta_1 x_t + \cdots + \beta_i x_{i,t} + \eta_t \quad (4)$$

with then η_t being the ARIMA process depicted in Eqs. (2) and (3). The resulting regression model thus is able to capture previous dynamics of a time series.

In harmonic regression models, the observed time series is considered as being composed of a signal, i.e., consisting of a sum of sinusoidal terms (or Fourier terms respectively) (Bloomfield 2000), so that any time series can be expressed as a combination of cosine (or sine) waves with differing periods. That is, the variation of a time series may be modeled as the sum of k different individual sinusoidal terms

(harmonics) occurring at different frequencies of periodic variation ω (Bloomfield 2000). Thus, a harmonic regression model can be defined as

$$y_t = \sum_{k=1}^K (\alpha_k \cos(\omega_k t) + \gamma_k \sin(\omega_k t)) + e_t \quad (5)$$

with white noise error e_t , coefficients α_k and γ_k , and $\omega_k, t = 1, 2, \dots, N$ being the frequencies of periodic variation.

Combining both the autocorrelated error term of (4) and the Fourier terms of (5) yields a DHR model (Young et al. 1999; Young 1999) using Fourier terms as predictors in combination with dynamic regression to handle periodic seasonality (Hyndman and Athanasopoulos 2018)

$$y_t = \beta_0 + \sum_{k=1}^K (\alpha_k \cos(\omega_k t) + \gamma_k \sin(\omega_k t)) + \eta_t \quad (6)$$

with η_t being a non-seasonal ARIMA (p,d,q) process. The DHR as in (6) and its extensions have already been utilized by research to forecast sub-daily call arrivals (Taylor 2008; Tych et al. 2002) since it allows for long seasonal periods compared to ARIMA and ETS models and short-term dynamics are handled by the ARIMA error (Hyndman and Athanasopoulos 2018). Nevertheless, the DHR model in (6) does not include additional contextual information such as the effect of holidays, catalog mailings, or billing cycles. As mentioned in the previous section, prior research (Aldor-Noiman et al. 2009; Andrews and Cunningham 1995) found such information to substantially enhance forecast accuracy and thus, recommended to include predictor variables (Taylor 2008). Therefore, we extend the DHR model in (6) from extant call center literature by adding predictor variables aside from Fourier terms:

$$y_t = \beta_0 + \beta_1 x_t + \dots + \beta_i x_{i,t} + \sum_{k=1}^K (\alpha_k \cos(\omega_k t) + \gamma_k \sin(\omega_k t)) + \eta_t \quad (7)$$

4 Analysis of customer support call arrival series

4.1 Preliminary data analysis

We gathered call center data from a leading German online retailer for fashion. The retailer's call center comprises arrival queues for each customer support, order taking, customer complaints, and consultation service. We investigate the customer support queue for e-mail arrivals as well as for call arrivals. Thus, in line with the concept of multi-channel customer contact in modern call centers, we are able to investigate two different queues of the same call center that represent different channels and types of media for customer support. The call queue is open from 7 a.m. to 10 p.m. from Monday through Saturday. Since our proposed DHR model is computationally only capable of modeling one seasonality, we apply a common two-step temporal aggregation

approach (Kourentzes et al. 2017): first, we aggregate our original high sampling frequency data (half-hourly data) at a pre-specified aggregation level with lower frequency (daily values) and thus, predict the daily arrival volume. Then, we disaggregate the daily predictions with respect to the averaged arrival distribution per weekday to re-yield the original high frequency, i.e., half-hourly predictions. The described temporal aggregation approach has gained substantial attention in recent methodological forecasting literature (Boylan and Babai 2016; Kourentzes et al. 2014; Kourentzes and Petropoulos 2016; Nikolopoulos et al. 2011) as it smooths the original time series, removes noise, improves forecast accuracy, and particularly simplifies the generation of forecasts (Kourentzes et al. 2017).

Hence, our aggregated daily data contain 1045 observations from January 2, 2016 to May 4, 2019, i.e., 174 weeks of data. One week comprises six observations and 1 year contains 312.25 observations considering leap years. The maximum number of call arrivals per day are 5300 arrivals and the mean of call arrivals per day is 2105.56. Our data exhibits overdispersion with a variance of 675,914.17. We excluded 2 weeks of data (i.e., 12 observations) due to incorrect interval capturing.

To enable the use of time series models, the time series has to be stationary. We conduct an Augmented Dickey Fuller (ADF) test (Dickey and Fuller 1979) to check for unit root in our data. With a p -value of 0.99 at lag order 312 (value of test statistic 0.1139), we cannot reject the null hypothesis of unit root in our data and thus, assume our data to be nonstationary.

To determine appropriate frequencies for our DHR model and further, to make the time series stationary, we have to assess the degree of seasonality in our data. Figure 1 depicts the overall daily call arrival volume of the customer support queue and its smoothed trend. The volume remains relatively constant with a slight increase towards the beginning of the year 2018 and a decrease towards the end of the dataset. Figure 2 displays three consecutive weeks of data. The number of call arrivals peaks on Mondays, then decreases throughout the week, and exhibits a second peak during the course of the week. The arrival volume drops substantially on Saturdays. As this pattern repeats every week, we assume daily seasonality, i.e., $s = 6$. We do not assume yearly seasonality as the yearly pattern is rather weak or non-existent respectively, considering Fig. 1. More formally, these findings further have been confirmed by the data's periodogram with a peak value at frequency $\omega = 0.1667$ and thus, the dominant period $T = \frac{1}{\omega}$ is 5.9999, i.e., it takes approximately 6 days to complete a full cycle. The periodogram does not indicate yearly seasonality.

We model different predictor variables for our data to improve forecast accuracy, summarized in Table 1. We utilize previous findings of literature regarding useful predictor variables (see e.g., Aldor-Noiman et al. 2009; Andrews and Cunningham 1995; Ibrahim et al. 2016) and include additional, potentially meaningful variables indicated by call center management or identified in preceding research (Rausch and Albrecht 2020).

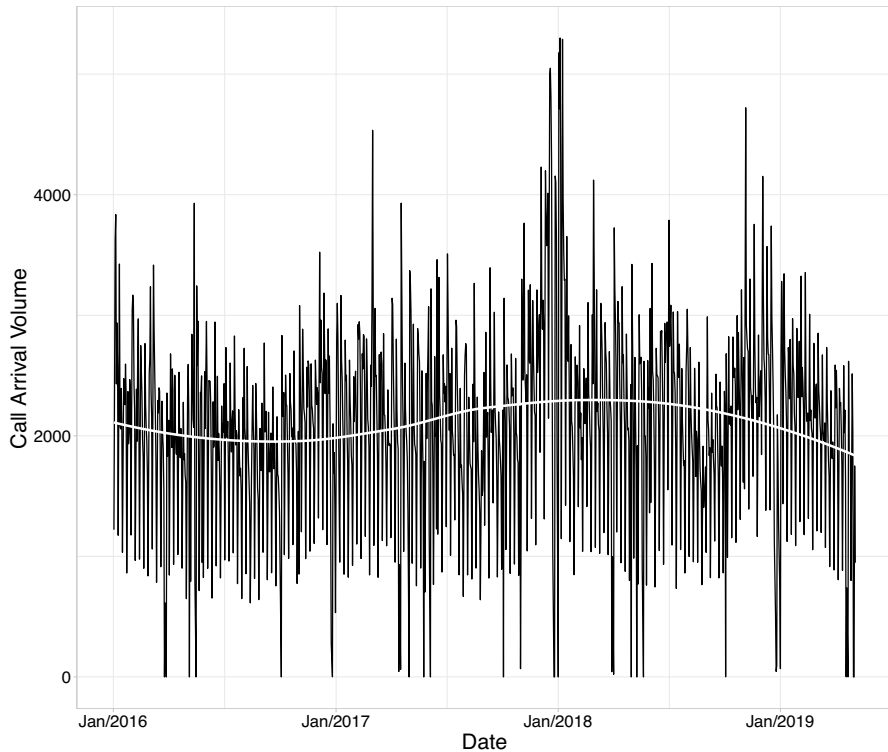


Fig. 1 Overall call arrival volume of customer support queue

4.2 Experimental design

To evaluate the performance of our proposed DHR model with predictor variables, we draw on several standard forecasting techniques of those presented in Sect. 2 for comparison listed in Table 2. Thereby, we utilize different time series models such as ARIMA, ETS, RW, TBATS, and further, the standard DHR approach without predictor variables as well as common high-performance ML methods such as RF and gradient boosting with L1 regularization (GBR) as benchmark approaches since they have been found to outperform other models by extant forecasting research. We do not include neural networks in our comparison due to the trade-off between model parsimony and forecast accuracy. Neural networks are frequently computationally infeasible in terms of computation time and complexity and thus, are rather unsuitable for practical use.

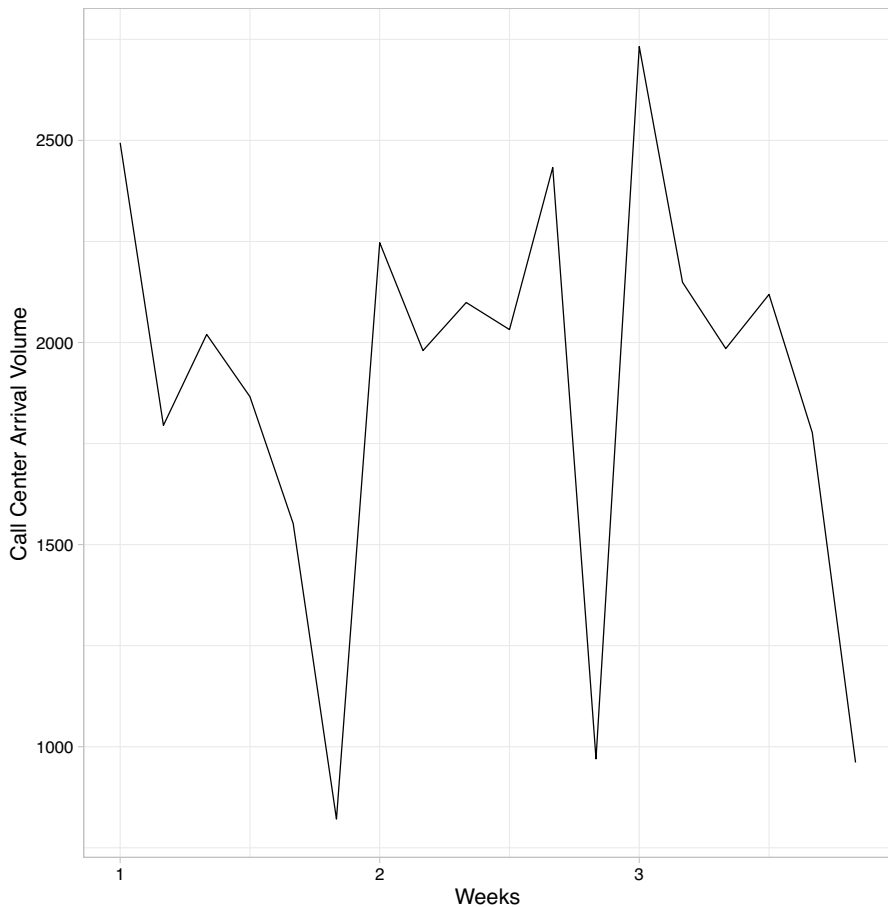


Fig. 2 Call arrival volume of three consecutive weeks

Since our time series is nonstationary, we apply time series decomposition before generating predictions with the time series approaches.² The seasonal-trend decomposition based on Loess (STL) (Cleveland et al. 1990) detrends and deseasonalizes the data yielding $y_t = \hat{S}_t + \hat{A}_t$ for additive decomposition ($y_t = \hat{S}_t * \hat{A}_t$ for multiplicative decomposition respectively) with $\hat{A}_t = \hat{T}_t + \hat{R}_t$ for additive decomposition ($\hat{A}_t = \hat{T}_t * \hat{R}_t$ for multiplicative decomposition respectively). Thereby, \hat{S}_t is the seasonal component and \hat{A}_t is the data without seasonality, i.e., the seasonally adjusted component. Both components are forecasted separately. The former is predicted by

² One of the anonymous reviewers pointed out that ETS models do not necessarily need time series decomposition prior to generating predictions. Nevertheless, although all ETS models are nonstationary, we also decompose the time series before applying the ETS model as predictions are assumed to become more accurate by detrending and deseasonalizing the time series first compared to predictions based on the global series Theodosiou (2011).

Table 1 Predictor variables

| Variable | Description | References |
|-------------------|--|--|
| Day-of-the-week | Nominal variable: six values (Monday to Saturday) to capture the day-of-the-week effect | Aldor-Noiman et al. (2009), Ibrahim et al. (2016), Ibrahim and L'Ecuier (2013) |
| Holiday | Nominal variable: 16 values (15 public holidays and ordinary weekdays) to capture the effect of German public holidays | Andrews and Cunningham (1995) |
| Day-after-holiday | Dummy variable: two values (days after public holidays and ordinary weekdays) to capture the effect of days after public holidays | |
| Outlier | Nominal variable: four values (extreme outliers and outliers (marked as such by management: if an extraordinary high number of arrivals can be explained by e.g. a special offer), days on which the call center is closed, and ordinary weekdays) to capture the effect of outliers | |
| School holidays | Metric variable indicating the number of German states having school holidays to capture the effect of German school holidays | |
| Year | Nominal variable: eight values (semiannual sections from January 2016 to May 2019) to capture the effect of busier seasons and the long-term development of the arrivals' level | |
| CW0-3 | Four dummy variables to capture the temporal effect of catalog mailings during the first weekend (CW0), the first week (CW1), the second week (CW2), and the third week (CW3) after release to capture the temporal effect of catalog mailings | Aldor-Noiman et al. (2009), Andrews and Cunningham (1995) |
| MMail1-2 | Six dummy variables to capture the temporal effect of e-mail reminders (MMail) as well as postal reminders (MPost) and due date e-mails (DMail) on the day of delivery (MMail1, MPost1, and DMail1) and the day after (MMail2, MPost2, and DMail2) | Aldor-Noiman et al. (2009) |

Table 2 Models for comparison

| Model | Description |
|-------------------------------|--|
| Time series models | |
| DHR | Combination of a dynamic regression model and a harmonic regression model, i.e., a regression model with Fourier terms as predictors and an ARIMA error term (Young et al. 1999; Young 1999) |
| STL + (non-seasonal) ARIMA | Combination of a seasonal-trend decomposition of time series based on Loess and a non-seasonal ARIMA (p,d,q) model, i.e., predictions are generated based on prior values y_{t-p} of y_t and prior errors ε_{t-q} (Box and Jenkins 1970) |
| STL + ETS | Combination of a seasonal-trend decomposition of time series based on Loess and an exponential smoothing innovation state space model (exponential smoothing model with an error term), i.e., predictions are the exponentially weighted average of past observations (Hyndman et al. 2002) |
| STL + RW with drift (RWDRIFT) | Combination of a seasonal-trend decomposition of time series based on Loess and a random walk model, i.e., predictions equal the last observation and the average of changes between consecutive observations |
| TBATS | Exponential smoothing innovation state space model (ETS model) with a Box-Cox transformation (stabilizes the time series' variance), Fourier terms (trigonometric expression of seasonality terms for complex seasonality as well as high frequency of seasonality; allow seasonality to change over time), and an ARMA (p,q) correction (De Livera et al. 2011) |
| ML approaches | |
| GBR | Ensemble of successive weak learners (i.e., models that achieve accuracy just above random guessing): within boosting, weak learners are trained sequentially trying to correct its respective predecessor (Schapire et al. 1998). I.e., each learner is constructed using feedback from previously grown learners. More specifically, within gradient boosting, a subclass of boosting, weak learners are fitted to the residual errors made by preceding learners and gradient descent is used to identify the errors in previous predictions (Friedman 2001, 2002) |
| RF | Ensemble of successive decision trees: Withing bagging, each learner is grown independently from earlier learners. I.e., each tree is built using a bootstrap sample of the data (Breiman 1996). More specifically, within random forests, a subclass of bagging, the algorithm draws n_{tree} bootstrap samples, grows an unpruned regression tree for each sample by randomly sampling m_{try} of the predictors at each node, and then chooses the best split among them. The outputs of the n_{tree} trees are aggregated and averaged to produce one final prediction (Breiman 2001; Liaw and Wiener 2002). An additional L1 regularization prevents the model from overfitting |

drawing on the last period of the estimated component, which equals a seasonal naïve method, whereas any non-seasonal forecasting approach can be utilized for the latter. The transformations of the decomposed time series are then inverted to yield the forecasts of the original time series $\{Y_t\}$ (Hyndman and Athanasopoulos 2018).

To select appropriate predictor variables and filter out uninformative ones for our proposed DHR model, we draw on the forward variable selection procedure (Hyndman and Athanasopoulos 2018). A prevailing approach to identify essential predictor variables is to drop those variables whose p -values are statistically insignificant (see e.g., Aldor-Noiman et al. 2009). However, in a forecasting context, the p -value does not necessarily determine the variable's predictive performance regarding the out-of-sample predictions which are practically relevant. Hence, we begin with the null model comprising none of the variables and add each predictor variable at a time. The variable is maintained if it enhances forecast accuracy. This step is repeated until no further improvement of accuracy is yielded.

We evaluate model performance with time series cross-validation and an expanding rolling window. We use 118 weeks of data comprising 709 observations, i.e., the observations from January 2, 2016 to April 7, 2018, as our initial training data. We fit the models and predict 1 week or six observations respectively (i.e., forecast horizon $h = 6$). During the next iteration, we roll the training data 1 week forward, re-estimate our models, and predict one unit of our forecast horizon further. During each iteration $n = 1, 2, \dots, N$, the ML models' hyperparameters are optimized by implementing tenfold cross-validation with grid search. We further optimize the number of Fourier terms k of the DHR models by including a second loop within each iteration n : Since k can have a maximum value of $T/2$, the grid search for k is set within the range $[1;3]$ and $k \in \mathbb{N}$. k is then optimized by fitting the model with the current training data and predicting the current out-of-sample data during each iteration $n = 1, 2, \dots, N$. Forecast accuracy for the out-of-sample predictions is calculated for each k and subsequently, k is chosen with respect to the highest forecast accuracy results. Overall, the procedure is repeated 52 times, i.e., for 1 year, and hence, $N = 52$. As stated earlier, we exclude 2 weeks of data from October 22, 2018 to November 4, 2018 due to incorrect interval capturing and therefore, we predict 300 daily call arrival volumes.

As stated earlier, to yield relevance for practical use, we further disaggregate our daily predictions with respect to the averaged call arrival distribution for each weekday per half-hour interval to invert the daily predictions back to the series' original half-hour frequency, see Fig. 3. We additionally include the averaged call distribution of holidays as they a divergent distribution compared to ordinary weekdays. Evidently, Mondays are the busiest days with a peak in the morning hours and a second smaller peak throughout the day. The remaining weekdays exhibit a similar course on a lower level. Saturdays register the fewest call arrivals on average throughout the day aside from holidays. Overall, each of the 300 predicted daily values is disaggregated into 30 half-hour intervals yielding a total of 9000 predictions.

To assess the models' performance, we compare the sub-daily forecasts with the out-of-sample or test data (i.e., the actual values) and compute forecast accuracy. We draw on mean absolute error (MAE) and root mean squared error (RMSE) as error measures:

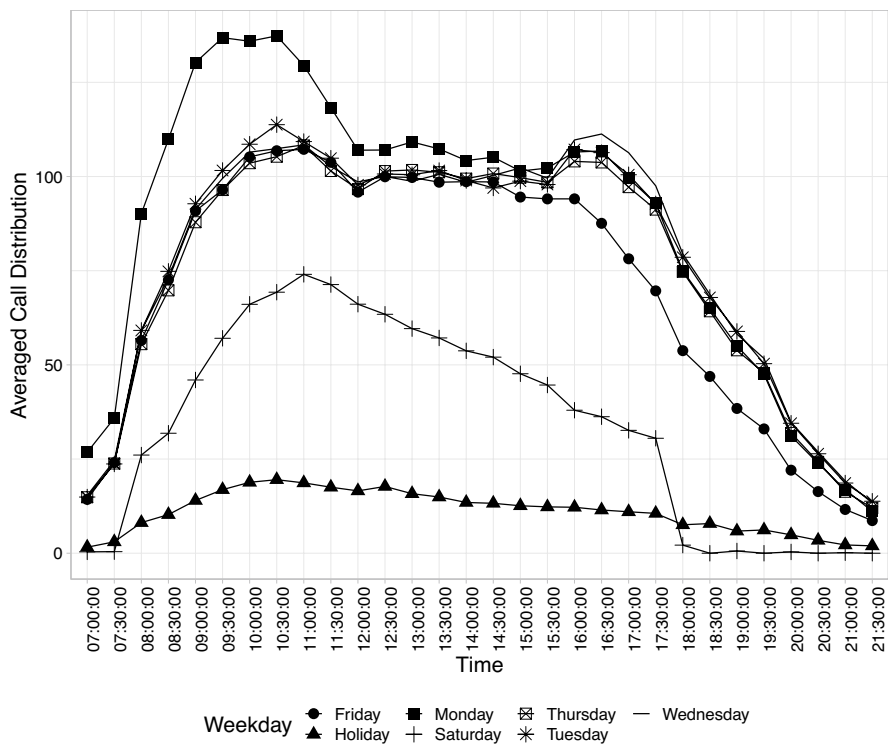


Fig. 3 Averaged call arrival distribution per weekday

$$\text{MAE} = \frac{1}{T} \sum_{i=1}^T |Y_i - \hat{Y}_i| \quad \text{RMSE} = \sqrt{\frac{1}{T} \sum_{i=1}^T (Y_i - \hat{Y}_i)^2}$$

with test (or out-of-sample) data Y_i , forecasted values \hat{Y}_i , and the number of forecasted values T . Since they are scale-dependent measures, they are both appropriate to compare predictions on the same scale. Both MAE and RMSE are frequently used by research to determine their forecasts' accuracy (see e.g., Aldor-Noiman et al. 2009; Barrow 2016; Ibrahim et al. 2016; Taylor 2008; Weinberg et al. 2007) as they can be calculated and interpreted easily (Hyndman and Athanasopoulos 2018).

We further check the robustness of our comparison results by considering (1) no lead time, (2) 1 week lead time, (3) 2 weeks lead time, and (4) 3 weeks lead time. In this context, lead time refers to the period between the dataset's last actual observation and the first created forecast.

Table 3 Forward variable selection MAE results for customer support call arrival forecasts

| | No lead time | One week | Two weeks | Three weeks |
|------------------------|----------------|----------------|----------------|----------------|
| No predictor variables | 18.6089 | 19.2526 | 19.7547 | 20.0912 |
| Day-of-the-week | 13.5712 | 13.6240 | 13.8336 | 14.1911 |
| Holiday | 11.7666 | 11.8471 | 11.9814 | 11.8789 |
| Day-after-holiday | 11.8434 | 11.9490 | 12.0462 | 12.0105 |
| School holidays | 11.8158 | 11.8982 | 12.0792 | 12.1299 |
| Outlier | 11.9839 | 12.0417 | 12.1656 | 12.1069 |
| Year | 12.1754 | 12.2806 | 12.4242 | 12.4873 |
| CW0 | 11.7389 | 11.8578 | 11.9731 | 11.8851 |
| CW1 | 11.7127 | 11.8451 | 11.9690 | 11.8648 |
| CW2 | 11.7655 | 11.8510 | 11.9737 | 11.8873 |
| CW3 | 11.7244 | 11.8144 | 11.9730 | 11.8720 |
| MMail1 | 11.7489 | 11.8264 | 11.9795 | 11.8644 |
| MMail2 | 11.5700 | 11.7269 | 11.9324 | 11.7686 |
| MPost1 | 11.3872 | 11.4660 | 11.6455 | 11.6049 |
| MPost2 | 11.3369 | 11.3959 | 11.5747 | 11.4863 |
| DMail1 | 11.3140 | 11.2669 | 11.5805 | 11.4648 |
| DMail2 | 11.2336 | 11.2250 | 11.5571 | 11.4688 |
| All | 11.5466 | 11.6480 | 11.9878 | 12.0333 |

The bold values are an improvement to the respective preceding value and the corresponding variable is included in the final model

4.3 Results

We conduct forward variable selection for our proposed DHR model. Drawing on the results in Table 3, the variables year, outlier, school holidays, day-after-holiday, and CW2 do not enhance forecast accuracy regardless of the considered lead time constellation whereas the variables day-of-the-week, holiday, CW1, MMail2, MPost1, MPost2, and DMail2 improve accuracy for every lead time. The highest MAE improvement is yielded by the day-of-the-week and the holiday variable indicating their importance for the call center arrival forecasts.

To gather more detailed insights on variable importance, we calculated the relative importance I of each predictor variable x

$$I_x = \frac{E^{(-x)} - E}{E}$$

with E being the MAE of the final model with all selected variables (i.e., the model selected via forward variable selection within Table 3) (Taieb and Hyndman 2014). The predictor variable increasing the error the most after removing its relative effect can be considered the most influential predictor variable given the other predictors. We calculate the variable importance with our proposed DHR model without lead time as this yielded the highest forecast accuracy among all considered lead times.

Table 4 Predictor variable importance for customer support call arrival forecasts

| | MAE without predictor variable (compared to selected model) | Predictor variable importance |
|-----------------|---|-------------------------------|
| Day-of-the-week | 14.6435 | 0.3035 |
| Holiday | 12.9488 | 0.1527 |
| CW0 | 11.2578 | 0.0022 |
| CW1 | 11.2812 | 0.0010 |
| MMail2 | 11.3541 | 0.0107 |
| MPost1 | 11.5830 | 0.0311 |
| MPost2 | 11.3290 | 0.0085 |
| DMail1 | 11.3356 | 0.0091 |
| DMail2 | 11.3083 | 0.0066 |

Table 5 MAE results for customer support call arrival forecasts

| | No lead time | One week | Two weeks | Three weeks |
|------------------------------|----------------|----------------|----------------|----------------|
| DHR with predictor variables | 11.2336 | 11.2250 | 11.5571 | 11.4648 |
| DHR | 14.3198 | 14.2243 | 14.3913 | 14.5242 |
| STL + ARIMA | 17.8254 | 17.3990 | 17.7603 | 18.2596 |
| STL + ETS | 17.8028 | 17.4140 | 17.8213 | 18.4370 |
| STL + RWDRIFT | 17.7657 | 17.5317 | 17.7757 | 18.2407 |
| TBATS | 16.3675 | 17.1001 | 17.2664 | 17.6892 |
| GBR | 13.4358 | 13.7363 | 13.7492 | 13.5698 |
| RF | 12.0249 | 12.0516 | 12.3048 | 12.3006 |

The highest forecast accuracy for each lead time is marked in bold

DHR: dynamic harmonic regression; GBR: gradient boosting with regularization; RF: Random forest; STL+ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL+ETS: innovation state space model with time series decomposition based on Loess; STL+RWDRIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box–Cox transformation, ARMA errors, trend and seasonal components

Table 4 depicts the resulting variable importances. The day-of-the-week as well as the holiday variable are clearly the two most important predictors. Among the variables containing information about marketing actions, the predictor capturing the effect of postal reminders on the day of delivery (MPost1) worsens forecast accuracy the most.

Tables 5 and 6 present the MAE and RMSE results of the investigated models. Our proposed DHR model with predictor variables outperforms the remaining models with respect to its MAE results for every considered lead time. Considering 2 weeks and 3 weeks of lead time, RF performs slightly better than our DHR model with regard to RMSE as evaluation metric. As the RMSE gives a higher weight to large errors, this indicates that our model made fewer large errors compared to RF. Nevertheless, as the discrepancy in RMSE between both models is

Table 6 RMSE results for customer support call arrival forecasts

| | No lead time | One week | Two weeks | Three weeks |
|------------------------------|----------------|----------------|----------------|----------------|
| DHR with predictor variables | 15.7521 | 16.6844 | 17.0088 | 17.1805 |
| DHR | 22.4746 | 22.9245 | 23.1713 | 23.4540 |
| STL + ARIMA | 30.3195 | 29.0067 | 29.3980 | 30.2608 |
| STL + ETS | 30.1632 | 28.9110 | 29.4276 | 30.4809 |
| STL + RWDRIFT | 30.1719 | 28.9619 | 29.1190 | 29.8225 |
| TBATS | 24.8822 | 26.0952 | 26.3939 | 27.0433 |
| GBR | 18.6523 | 19.2414 | 19.0341 | 18.7706 |
| RF | 16.6278 | 16.7476 | 16.9844 | 16.9797 |

The highest forecast accuracy for each lead time is marked in bold

DHR: dynamic harmonic regression; GBR: gradient boosting with regularization; RF: Random forest; STL+ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL+ETS: innovation state space model with time series decomposition based on Loess; STL+RWDRIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box–Cox transformation, ARMA errors, trend and seasonal components

only around 0.2, this seems rather negligible. This further indicates the superiority and importance of contextual information.

The decomposed ARIMA, ETS, and RW models yield the most inaccurate forecasts and among the time series models while the standard DHR model without predictor variables is the second best performing model. Comparing both DHR models, predictor variables apparently enhance forecast accuracy substantially. Further, since the ML models comprise predictor variables, their forecasts are comparable but slightly worse than those of our proposed model. Overall, ML models are superior to the time series models.

Evidently, ex-post forecasts (i.e., their predictor variables can be modeled for both past observations (the training data) as well as future observations (the out-of-sample data)) of our proposed model and the ML models are outperforming ex-ante forecasts (i.e., the models are only using information that is available at the time of generating the forecasts) of the time series models used.

Considering the lead times, the present results support previous findings in call center forecasting literature (Ibrahim et al. 2016; Rausch and Albrecht 2020). Forecast accuracy declines steadily with increasing lead time for most of the models and the most accurate predictions for each model are yielded without any lead time.

4.4 Robustness checks

To test the robustness of our results, we conducted further analyses. Although the temporal aggregation approach is assumed to remove noise and enhance forecast accuracy, it would not have been mandatory for the benchmark time series models, as these are computationally capable of modeling more than one seasonality. Thus, we generated forecasts based on the original series (i.e., sub-daily data) without the temporal aggregation approach (see Tables 7 and 8). Although forecast accuracy partially improves, our proposed DHR model with predictor variables is still

Table 7 MAE results for customer support call arrival forecasts based on original sub-daily data without temporal aggregation

| | No lead time | One week | Two weeks | Three weeks |
|---------------|--------------|----------|-----------|-------------|
| DSHW | 68.9264 | 151.4902 | 230.738 | 310.4899 |
| STL + ARIMA | 14.5263 | 14.7404 | 15.5448 | 15.8520 |
| STL + ETS | 14.5263 | 14.7407 | 15.5448 | 15.8520 |
| STL + RWDRIFT | 14.6651 | 14.6334 | 15.2941 | 15.7877 |
| TBATS | 18.1162 | 17.9856 | 18.8661 | 18.9823 |
| GBR | 12.9393 | 13.1488 | 13.3987 | 13.7386 |
| RF | 11.7544 | 11.8129 | 12.0648 | 12.8134 |

DSHW: double seasonal Holt-Winters; GBR: gradient boosting with regularization; RF: Random forest; STL + ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL + ETS: innovation state space model with time series decomposition based on Loess; STL + RWDRIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box–Cox transformation, ARMA errors, trend and seasonal components

Table 8 RMSE results for customer support call arrival forecasts based on original sub-daily data without temporal aggregation

| | No lead time | One week | Two weeks | Three weeks |
|---------------|--------------|-----------|-----------|-------------|
| DSHW | 880.8499 | 1,769.499 | 2,675.125 | 3,584.931 |
| STL + ARIMA | 22.7009 | 23.1810 | 24.2187 | 25.0726 |
| STL + ETS | 22.9251 | 23.0876 | 23.9239 | 24.7768 |
| STL + RWDRIFT | 23.0503 | 23.1555 | 23.9506 | 24.7793 |
| TBATS | 30.8184 | 30.7156 | 31.6453 | 31.8839 |
| GBR | 18.1216 | 18.3299 | 18.6043 | 19.3079 |
| RF | 15.5678 | 16.6541 | 16.8929 | 18.4903 |

DSHW: double seasonal Holt-Winters; GBR: gradient boosting with regularization; RF: Random forest; STL + ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL + ETS: innovation state space model with time series decomposition based on Loess; STL + RWDRIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box–Cox transformation, ARMA errors, trend and seasonal components

superior in terms of MAE results. Nevertheless, without temporal aggregation, RF's RMSE values for forecasts without lead time and with one as well as 2 weeks lead time are slightly better than those of the DHR model. Additionally, we generated forecasts based on the original sub-daily data with a double seasonal exponential smoothing model (DSHW) and found our models to be noticeably superior.

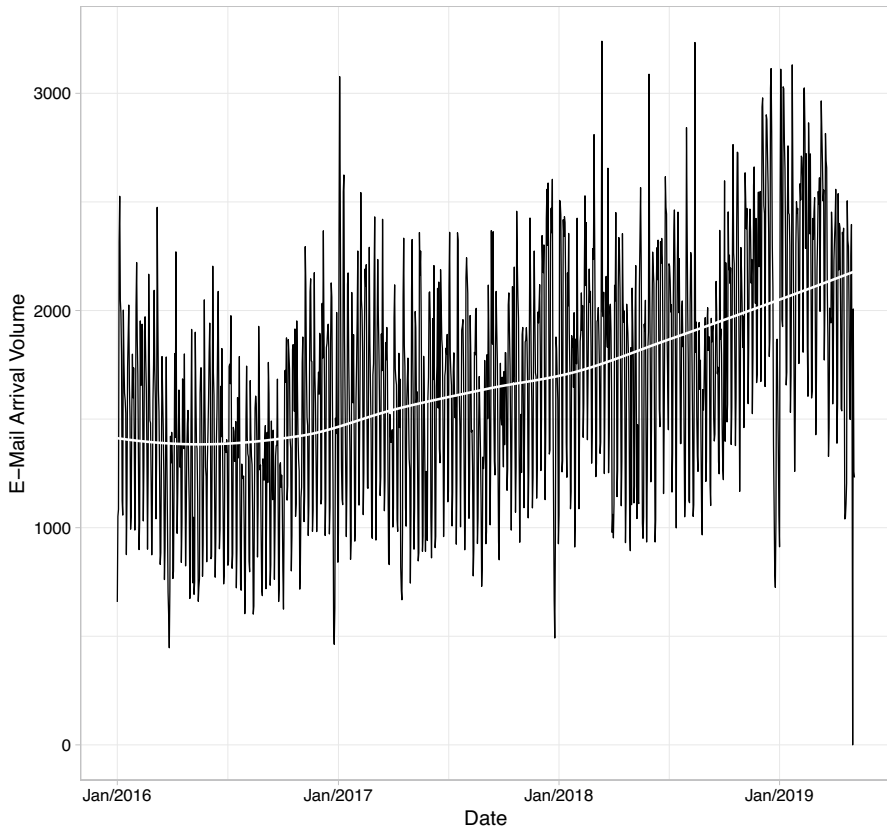


Fig. 4 Overall e-mail arrival volume of customer support queue

5 Analysis of customer support e-mail arrival series

5.1 Preliminary data analysis

To further validate our results and allow for the development of multi-channel customer contact in modern call centers, we additionally investigate the e-mail arrivals of the customer support queue. The incoming e-mail data of this queue varies from the previous call analysis in the number of available observations per week, the level of average and maximum arrival count per interval, and the existence of trend in the arrival volume.

The aggregated daily data consist of 1220 observations from January 2, 2016 to May 5, 2019, i.e., 174 weeks of data. Since e-mails arrive at any time throughout the day and on every weekday from Mondays to Sundays, 1 week consists of seven observations and 1 year comprises 365.25 observations considering leap years. The maximum number of e-mail arrivals is 3240 per day. The customer support queue receives 1670.15 e-mails on average each day and with a variance of 283,687.87 the

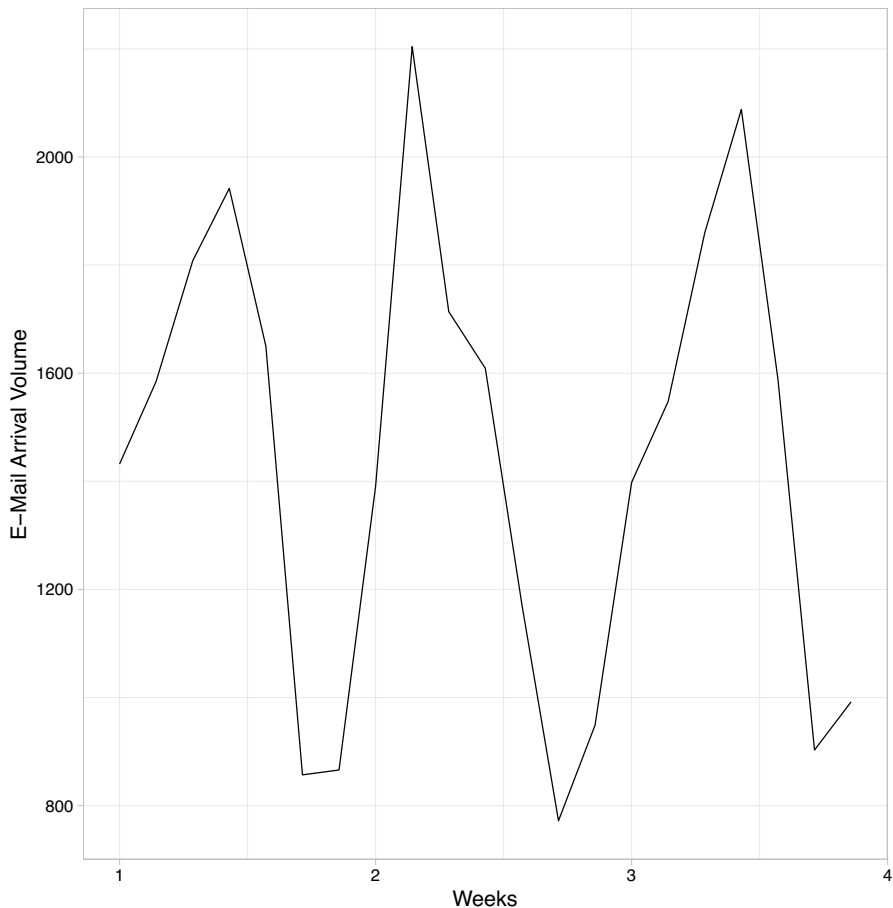


Fig. 5 E-Mail arrival volume of three consecutive weeks

data is overdispersed. With a p -value of 0.98 at lag order 365 (value of test statistic -0.4533) of the ADF test, we cannot reject the null hypothesis of unit root in our data and thus, assume our data to be nonstationary.

Figure 4 exhibits a steady upwards trend in the overall e-mail arrival volume. Figure 5 reveals that the e-mail arrival volume is high from Mondays to Fridays and drops noticeably on weekends. Since this pattern repeats every week, we assume weekly seasonality for the daily data, i.e., $s = 7$. Further, there is no considerable yearly seasonality. More formally, the data's periodogram exhibits a peak value at frequency $\omega = .1432$ and thus, the dominant period $T = \frac{1}{\omega}$ is 6.9832, i.e., it takes approximately seven days to complete a full cycle. The periodogram does not indicate a frequency which implies yearly seasonality.

We use the experimental design and predictor variables described in Sects. 4.1 and 4.2 respectively. Accordingly, we extend the day-of-the-week predictor variable as well as the forecast horizon h to 7 days. To determine the averaged e-mail

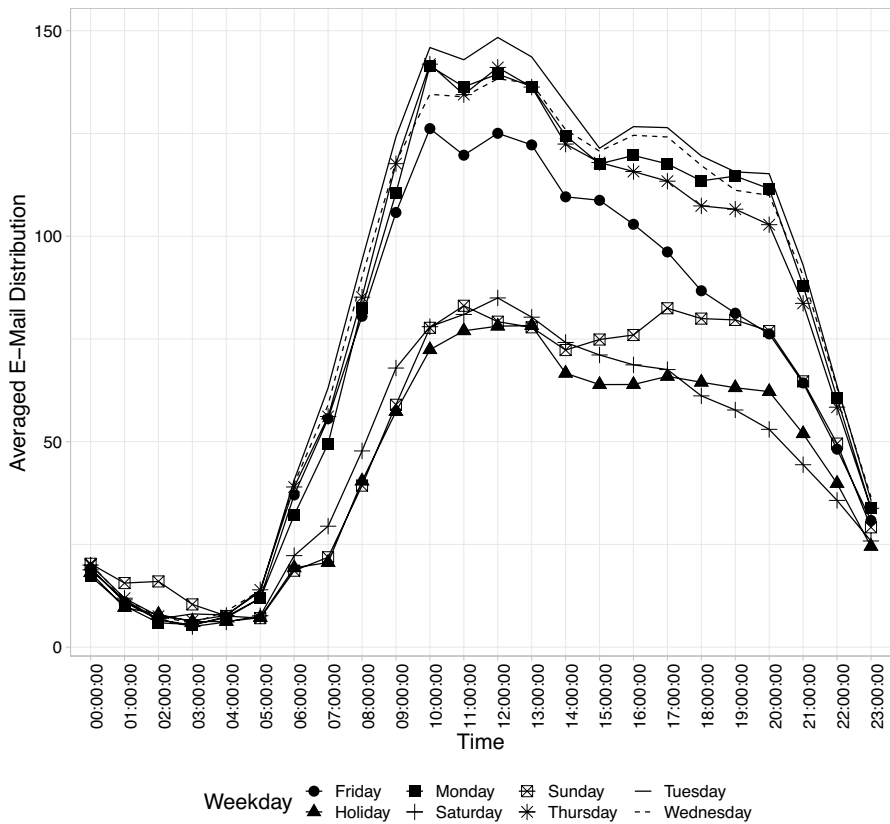


Fig. 6 Averaged e-mail arrival distribution per weekday

distribution per weekday, we draw on the original hourly dataset comprising 29,280 observations and average the cumulated e-mail arrivals for each interval per weekday (see Fig. 6). Evidently, Mondays to Fridays exhibit a similar distribution with few e-mail arrivals during the night, a peak in the morning hours, and a second smaller peak throughout the day. The volume drops towards the end of the day. On Saturdays, Sundays, and holidays the e-mail arrival volume is relatively low but has a similar distribution.

5.2 Results

Similar to the customer support call arrival series, we conduct forward variable selection for our DHR model listed in Table 9. The variables day-after-holiday, school holidays, outlier, year, CW1, CW2, and DMail1 do not improve forecast accuracy regardless of the considered lead time constellation, whereas the variables day-of-the-week, holidays, CW0, and MMail1 have a positive impact on accuracy for every lead time. Analogous to the call arrival series,

Table 9 Forward variable selection MAE results for customer support e-mail arrival forecasts

| | No lead time | One week | Two weeks | Three weeks |
|------------------------|----------------|----------------|----------------|----------------|
| No predictor variables | 16.4636 | 17.3453 | 18.1969 | 18.5699 |
| Day-of-the-week | 14.6233 | 14.7007 | 15.0112 | 15.1020 |
| Holiday | 14.2472 | 14.3957 | 14.5968 | 14.5782 |
| Day-after-holiday | 14.2547 | 14.4004 | 14.5999 | 14.5856 |
| School holidays | 14.2520 | 14.4547 | 14.6330 | 14.6229 |
| Outlier | 14.2759 | 14.4185 | 14.5990 | 14.6019 |
| Year | 14.3672 | 14.5808 | 14.8253 | 14.8272 |
| CW0 | 14.2470 | 14.3674 | 14.5592 | 14.5604 |
| CW1 | 14.2881 | 14.3853 | 14.5636 | 14.5761 |
| CW2 | 14.2483 | 14.3826 | 14.5679 | 14.5792 |
| CW3 | 14.2476 | 14.3685 | 14.5588 | 14.5686 |
| MMail1 | 14.0404 | 14.2270 | 14.4386 | 14.4538 |
| MMail2 | 14.0306 | 14.2055 | 14.4413 | 14.4627 |
| MPost1 | 14.0497 | 14.2740 | 14.4738 | 14.4513 |
| MPost2 | 14.0283 | 14.2080 | 14.4336 | 14.4520 |
| DMail1 | 14.0533 | 14.3239 | 14.5907 | 14.5382 |
| DMail2 | 13.9974 | 14.1827 | 14.4355 | 14.4230 |
| All | 14.2046 | 14.4940 | 14.8711 | 14.9533 |

The bold values are an improvement to the respective preceding value and the corresponding variable is included in the final model

Table 10 Predictor variable importance for customer support e-mail arrival forecasts

| | MAE without predictor variable (compared to selected model) | Predictor variable importance |
|-----------------|---|-------------------------------|
| Day-of-the-week | 15.6057 | 0.1149 |
| Holiday | 14.3916 | 0.0282 |
| CW0 | 14.0072 | 0.007 |
| MMail1 | 14.2565 | 0.0185 |
| MMail2 | 14.0121 | 0.0010 |
| MPost2 | 13.9975 | 0.0000 |
| DMail2 | 14.0283 | 0.0022 |

the day-of-the-week and holiday variables cause the highest MAE improvement and are assumed to be important elements contributing to the present prediction results.

To gather more detailed insights into predictor variable importance, we determine—analogously to the customer support call arrival series—the variable importance of each predictor included in the selected model from Table 9. Table 10 outlines the respective variable importance. Again, the day-of-the-week variable is the most important predictor. With respect to the marketing-relevant

Table 11 MAE results for customer support e-mail arrival forecasts

| | No lead time | One week | Two weeks | Three weeks |
|------------------------------|----------------|----------------|----------------|----------------|
| DHR with predictor variables | 13.9974 | 14.1827 | 14.4336 | 14.4230 |
| DHR | 14.9317 | 15.2102 | 15.4306 | 15.5281 |
| STL + ARIMA | 15.5983 | 15.6356 | 15.8708 | 15.9973 |
| STL + ETS | 15.7058 | 15.7584 | 15.9991 | 16.1562 |
| STL + RWDRIFT | 17.1299 | 17.6039 | 17.6647 | 17.9944 |
| TBATS | 15.6257 | 15.9940 | 16.2795 | 16.3678 |
| GBR | 16.6198 | 18.5997 | 17.3449 | 17.1729 |
| RF | 15.9811 | 16.6069 | 17.0371 | 17.1368 |

The highest forecast accuracy for each lead time is marked in bold

DHR: dynamic harmonic regression; GBR: gradient boosting with regularization; RF: Random forest; STL+ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL+ETS: innovation state space model with time series decomposition based on Loess; STL+RWDRIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box–Cox transformation, ARMA errors, trend and seasonal components

Table 12 RMSE results for customer support e-mail arrival forecasts

| | No lead time | One week | Two weeks | Three weeks |
|------------------------------|----------------|----------------|----------------|----------------|
| DHR with predictor variables | 25.8450 | 26.3258 | 26.7295 | 26.2177 |
| DHR | 27.2817 | 27.7663 | 27.9761 | 28.8035 |
| STL + ARIMA | 28.3968 | 28.4400 | 28.7987 | 29.0450 |
| STL + ETS | 28.4915 | 28.5459 | 28.9247 | 29.2177 |
| STL + RWDRIFT | 30.2348 | 31.1480 | 31.4703 | 31.5576 |
| TBATS | 28.5707 | 29.1941 | 29.4854 | 29.5885 |
| GBR | 28.9195 | 32.0761 | 29.9215 | 29.6442 |
| RF | 28.1436 | 28.8698 | 29.3342 | 29.4563 |

The highest forecast accuracy for each lead time is marked in bold

DHR: dynamic harmonic regression; GBR: gradient boosting with regularization; RF: Random forest; STL+ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL+ETS: innovation state space model with time series decomposition based on Loess; STL+RWDRIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box–Cox transformation, ARMA errors, trend and seasonal components

variables, the predictor capturing the effect of e-mail reminders on the day of delivery (MMail1) worsens forecast accuracy the most.

Tables 11 and 12 summarize the MAE and RMSE results for the customer support e-mail arrival forecasts. Although the customer support queue receives fewer e-mails than calls on average, the MAE and RMSE obtained for the customer support e-mail arrival predictions are comparable to those of the customer support call arrival forecasts. Similar to the call arrivals analysis, the proposed DHR model with predictor variables outperforms the remaining time series and ML approaches for every considered lead time constellation. Forecast accuracy

Table 13 MAE results for customer support e-mail arrival forecasts based on original sub-daily data without temporal aggregation

| | No lead time | One week | Two weeks | Three weeks |
|----------------|--------------|----------|-----------|-------------|
| DSHW | 18.4400 | 19.9364 | 20.5614 | 20.3284 |
| STL + ARIMA | 15.8875 | 17.0814 | 17.8252 | 17.8173 |
| STL + ETS | 17.1596 | 18.6449 | 19.1254 | 19.5232 |
| STL + RWDRIIFT | 17.7435 | 19.1402 | 19.6219 | 20.1637 |
| TBATS | 19.8545 | 20.8313 | 21.7826 | 21.9521 |
| GBR | 18.5135 | 19.0513 | 17.5940 | 17.4820 |
| RF | 15.5333 | 16.1990 | 16.8110 | 16.9005 |

DSHW: double seasonal Holt-Winters; GBR: gradient boosting with regularization; RF: Random forest; STL + ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL + ETS: innovation state space model with time series decomposition based on Loess; STL + RWDRIIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box-Cox transformation, ARMA errors, trend and seasonal components

Table 14 RMSE results for customer support e-mail arrival forecasts based on original sub-daily data without temporal aggregation

| | No lead time | One week | Two weeks | Three weeks |
|----------------|--------------|----------|-----------|-------------|
| DSHW | 32.9748 | 35.1945 | 35.9617 | 35.6316 |
| STL + ARIMA | 28.5696 | 30.0853 | 31.0859 | 30.7144 |
| STL + ETS | 29.4874 | 31.4311 | 32.1550 | 32.3495 |
| STL + RWDRIIFT | 20.2880 | 31.9048 | 33.1407 | 33.4434 |
| TBATS | 32.0302 | 33.2508 | 34.2860 | 34.3550 |
| GBR | 32.4089 | 29.5703 | 30.2263 | 30.3050 |
| RF | 27.7846 | 28.7869 | 29.2917 | 29.3207 |

DSHW: double seasonal Holt-Winters; GBR: Gradient boosting with regularization; RF: Random forest; STL + ARIMA: autoregressive integrated moving average with time series decomposition based on Loess; STL + ETS: Innovation state space model with time series decomposition based on Loess; STL + RWDRIIFT: random walk with drift with time series decomposition based on Loess; TBATS: trigonometric seasonality, Box-Cox transformation, ARMA errors, trend and seasonal components

mainly declines with higher lead times while the best performance for each model is yielded without lead time.

For the customer support e-mail arrivals, the performance gap between time series and ML models is not as evident as for the call arrivals. Under the conditions of a more apparent trend in the examined data, time series models operate in a comparable performance range as ML models. The DHR model without predictor variables yields the second-best forecasts indicating that the Fourier terms itself have a high predictive potential. The additional predictor variables in our model further enhance this predictive power. Among the ML models, the RF algorithm outperforms the gradient boosting approach with L1 regularization.

5.3 Robustness checks

We conducted additional analyses to test the results' robustness. We generated predictions based on the original sub-daily data without the temporal aggregation approach (see Tables 13 and 14). Forecast accuracy does not improve for the time series models, but for RF. Similar to the customer support call arrivals, we additionally generated forecasts based on the original sub-daily data with a double seasonal exponential smoothing model. Overall, our proposed DHR approach with predictor variables still outperforms the benchmark models in terms of both MAE as well as RMSE and for every lead time constellation.

6 Discussion

The findings of the analysis using call as well as e-mail arrival data from the customer support queue of the call center demonstrate clear benefits of the use of our proposed model. In line with the hypothesis that combining the strengths of different forecasting model types will lead to an increase in prediction performance and, at the same time, entail advantages for the use in practice, the DHR model with predictor variables outperforms other approaches investigated. Thereby, we contribute not only to the existing body of literature in several ways but further provide practical implications for decision-makers regarding methodological aspects on the one hand and meaningful contextual predictor variables on the other hand.

First, the results on both data sets show that our proposed DHR model with predictor variables yields better forecast accuracy than traditional time series models and ML approaches. Precisely, it outperforms established time series models used in previous research (e.g., Andrews and Cunningham 1995; Bianchi et al. 1998; De Livera et al. 2011; Hyndman et al. 2002) such as ARIMA, ETS, TBATS, standard DHR, and RW as well as powerful ML approaches such as gradient boosting and RF for every considered lead time constellation. This is achieved by simultaneously capturing the dynamics of the time series and including additional predictor variables. Previous studies on call arrivals forecasting methods only focused on of these capabilities in the same model. Thus, the standard DHR model as applied by extant literature (Taylor 2008; Tych et al. 2002) only relies on Fourier terms assuming that any time series can be expressed as a combination of cosine (or sine) waves with differing periods and on an ARIMA error term capturing short-term dynamics. At the same time, prior research suggests certain predictor variables to enhance forecast accuracy (Aldor-Noiman et al. 2009; Andrews and Cunningham 1995). We therefore specifically contribute to call center forecasting literature by bringing methodological strings of research together and, in doing so, substantially increase the accuracy of call arrival forecasts. Additional robustness and generalizability are added to the presented results by replicating them for two different series with distinctions in trend, number of observations, as well as level of average arrival count. Reflecting our findings in a more conceptual and abstract manner, we thus contribute to literature by finding evidence that such hybrid models (combining both time series models as well as models with contextual information) unveil a high predictive potential.

Drawing on a broader perspective regarding data characteristics, our results additionally suggest that the magnitude of trend in the time series should be considered in model selection. For data exhibiting only a slight trend like the call arrival series, ML models are outperforming traditional time series models. However, the latter are more competitive and comparable to ML models if the data has a stronger trend like our e-mail arrival series. Further, from a more general perspective of model selection, we found ex-post forecasts of models with predictor variables, i.e., the ML models and our proposed DHR model with predictor variables, to be predominantly more accurate than ex-ante forecasts of models without predictor variables, i.e., time series models, aligning with prior findings (Rausch and Albrecht 2020). This suggests that the general type of forecasting approach and its possibility of including contextual factors in the form of predictor variables particularly affects prediction accuracy in a practical call arrival forecast setting. Overall, preliminary call center forecasting literature recognized the predictive potential of ML approaches (Albrecht et al. 2021; Barrow 2016; Jalal et al. 2016; Rausch and Albrecht 2020) but is still in its infancy and thus, we substantiate the knowledge on the performance of ML models.

As call center managers strongly rely on the accuracy of call arrival predictions for staffing, the improvements achieved by our proposed model implicate high relevance for practice. To keep operating costs at a minimum by avoiding overstaffing and, at the same time, to maximize perceived service quality by shortening long waiting times caused by understaffing, the correct number of call center agents is crucial. Considering e.g. the customer support call arrivals' predictions without lead time of our proposed DHR model compared to ARIMA (most inaccurate model), call center managers would need approximately 4.12³ call center agents on average less per day in case the model overestimates the arrival volume. Accordingly, on average customers would need to wait approximately 0.93⁴ min less if the model underestimates the arrival volume. In comparison with the standard DHR model, the difference amounts to 1.93 agents less per day or respectively, an extension of customer waiting time of 0.43 min.

Additionally, results for both data sets fit with the theory that call center managers are recommended to minimize lead time in arrivals' forecasts, aligning with prior research (Ibrahim et al. 2016; Rausch and Albrecht 2020). For every model investigated, the best performance is yielded without lead time and generally forecast accuracy decreases steadily with longer lead times. However, in practice longer lead times are frequently mandatory due to personnel planning restrictions. Thus, to overcome this obstacle, managers might consider a two-stage forecasting process: first, producing an early forecast for the agents' scheduling with a pre-defined number of weeks in advance and then, adjusting this forecast right before the start of the predicted week. Our results indicate that the latter prediction with a shorter lead

³ If the processing time is 10 min per call arrival and the working hours per call center agent are 8 h per day.

⁴ If the processing time is 10 min per call arrival and there are 70.95 call arrivals per interval on average.

time is more accurate so that managers get more reliable information to incorporate immediate changes into the schedule.

Second, we increase existing knowledge on useful predictor variables. We determined the predictor variables' practical value by conducting a forward variable selection procedure and by calculating their variable importance. The results indicate that modeling the day of the week and holidays as predictor variables yields the highest improvement of forecast accuracy and thus, confirms prior research suggesting these influential factors (Aldor-Noiman et al. 2009; Andrews and Cunningham 1995; Brown et al. 2005; Ibrahim et al. 2016; Ibrahim and L'Ecuyer 2013). More formally, both variables were found to be the most important predictors, i.e., they worsen forecast accuracy the most in case they would be removed from the selected model. Moreover, the results illustrate that capturing the impact of catalog mailings during the first weekend (i.e., CW0) (and the first week after release (i.e., CW1) respectively) enhances prediction accuracy across all considered lead times for the customer support e-mail (and call arrival series respectively). This particularly indicates that variables including information on marketing actions such as mailings affect customer behavior in terms of e-mail and call volume directly after release. Further, it is shown that postal reminders on the day of their delivery (MPost1) are the most important marketing-related predictor variables. Thus, the results suggest that postal reminders have a substantial effect on the call arrival volume. Vice versa, reminders via e-mail affect the e-mail arrival volume on the day of delivery (MMail1) as this variable were found to be the most important marketing-related predictor for the e-mail arrival series. These findings extend existing literature since the effect of periods with catalog mailings or billing cycles has not been investigated over time (Aldor-Noiman et al. 2009; Andrews and Cunningham 1995). Thus, by partitioning billing and marketing mailing periods into sequential shorter periods, we are able to capture their temporal effect regarding the first weekend and the first, second, and third week after a catalog release as well as the day of a reminder's delivery and the day after.

Consequently, the results first encourage practitioners to include the availability of explanatory data in their considerations when selecting forecasting models in a call center context. Then, when choosing to use ex-post forecasting models, not only date dependent predictor variables as commonly suggested by literature (such as weekday and holidays) but also factors related to enterprises' customer contact activities need to be considered. Thus, on the one hand forecast accuracy for staffing and ergo high service quality can be improved while on the other hand valuable insights on the effect of activities such as mailings on customer behavior can be gained.

The theoretical and practical implications notwithstanding, our research is subject to limitations that stimulate future research. Noticeably, the proposed method is only capable of modeling one seasonality which can limit its use for complex data with multiple seasonality like e.g. sub-daily (i.e., half-hourly or hourly) data. We overcome this constraint by applying a two-step temporal aggregation procedure to yield sub-daily forecasts. Although this aggregation-disaggregation approach is common practice in forecasting literature, it nevertheless poses an additional obstacle compared to direct sub-daily predictions. Also, we did not test for the optimal

aggregation level for both series, i.e., whether to conduct temporal aggregation at a single level or at multiple levels. Besides, it is beyond the scope of this study to investigate different forecast horizons. To increase the reliability of results, in addition to considering different data sets and lead times, future research on this aspect of call center arrival forecasting is encouraged. Regarding model comparison, our study is limited to a selected range of commonly used models on the one hand and promising approaches from related fields on the other hand. In the authors' opinion, comparing the proposed DHR approach to more models, e.g., additional ML methods or mixed-effects models, as well as expanding its application to different businesses' call center data beyond online retail is a fruitful path for future research. In this connection, the high dependency of model performance on the availability and quality of explanatory data needs to be considered. Furthermore, future research can validate our findings on such hybrid models and confirm their superiority by combining the strengths of different model types.

7 Conclusion

Call centers constitute an important customer touchpoint for many businesses. To achieve a high level of customer service satisfaction through short waiting times and good customer support, providing the appropriate number of agents is a critical task for call center management. For this purpose, accurate and feasible forecasting methods to predict call center arrival volumes are needed.

Combining the strengths of different model types investigated by previous research, this study proposes a new method for call center arrivals' forecasting that is able to capture the dynamics of time series and, at the same time, include contextual information in the form of predictor variables. We hypothesize that this approach leads to an increase in prediction performance while also yielding advantages for practical use. The implemented forecasting method extends the established DHR model, which utilizes a sum of sinusoidal terms as predictors to handle periodic seasonality and an ARIMA error to capture short-term dynamics, by including predictor variables in the considered information space to generate predictions. To test the predictive potential of our approach, we analyzed two datasets comprising 174 weeks of data on the call and e-mail arrivals of the customer support queue of a leading German online retailer. We compare our method to traditional time series models (i.e., ARIMA, ETS, TBATS, and RW) as well as established ML approaches (i.e., RF and GBR). Further, we apply time series cross-validation and an expanding rolling window over 52 weeks to assess model performance.

Results show that our proposed DHR model with predictor variables outperforms traditional time series models and ML approaches with regard to forecast accuracy for both data sets and in all lead time constellations investigated. Reflecting this on a more abstract level, we find evidence that such hybrid models combining the benefits of both model types unleash a high predictive potential. Moreover, for the chosen ex-post forecasting method, the predictor variables' practical value can be determined by conducting a forward variable selection procedure. Beyond confirming date-related variables (such as weekday and holidays) as important influential

factors for arrival volume, it is shown that catalog mailings and billing cycles exhibit a periodically enhancing effect on prediction accuracy when implemented as predictor variables over time. With the present study we contribute to existing literature by developing a new powerful method to be used in call center arrival forecasting as well as adding knowledge on the temporal effect of predictor variables on customer call and e-mail behavior in this context. We showed that data on e-mail reminders are particularly helpful to predict e-mail arrivals and vice versa, postal reminders are helpful to predict call arrivals. The model's ability to capture both time series information and predictor variables is well suited for the dynamic environment of practical call center arrival forecasting as it not only provides robust forecasts but also offers valuable insights into the effect of a company's customer contact activities. In this regard, future research on the use of other hybrid models in call center arrival forecasting is encouraged to broaden the spectrum of highly accurate methods with practical relevance.

Funding Open Access funding enabled and organized by Projekt DEAL. This research received funding from the Bayerisches Staatsministerium für Wirtschaft, Landesentwicklung und Energie (StMWi).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aksin OZ, Armony M, Mehrotra V (2007) The modern call center: a multi-disciplinary perspective on operations management research. *Prod Oper Manag* 16:665–688
- Albrecht T, Rausch TM, Derra ND (2021) Call me maybe: methods and practical implementation of artificial intelligence in call center arrivals' forecasting. *J Bus Res* 123:267–278. <https://doi.org/10.1016/j.jbusres.2020.09.033>
- Aldor-Noiman S, Feigin PD, Mandelbaum A (2009) Workload forecasting for a call center: methodology and a case study. *Ann Appl Stat* 3:1403–1447. <https://doi.org/10.1214/09-AOAS255>
- Andrews BH, Cunningham SM (1995) L. L. Bean improves call-center forecasting. *Interfaces* 25:1–13
- Avramidis AN, Deslauriers A, L'Ecuyer P (2004) Modeling daily arrivals to a telephone call center. *Manag Sci* 50:896–908. <https://doi.org/10.1287/mnsc.1040.0236>
- Barrow DK (2016) Forecasting intraday call arrivals using the seasonal moving average method. *J Bus Res* 69:6088–6096. <https://doi.org/10.1016/j.jbusres.2016.06.016>
- Barrow DK, Kourentzes N (2018) The impact of special days in call arrivals forecasting: a neural network approach to modelling special days. *Eur J Oper Res* 264:967–977. <https://doi.org/10.1016/j.ejor.2016.07.015>

- Bianchi L, Jarrett J, Choudary Hanumara R (1998) Improving forecasting for telemarketing centers by ARIMA modeling with intervention. *Int J Forecast* 14:497–504. [https://doi.org/10.1016/s0169-2070\(98\)00037-5](https://doi.org/10.1016/s0169-2070(98)00037-5)
- Bloomfield P (2000) *Fourier analysis of time series: an introduction*, 2nd edn. Wiley, New York
- Box GEP, Jenkins GM (1970) *Time series analysis: forecasting and control*. Holden-Day, San Francisco
- Boylan JE, Babai MZ (2016) On the performance of overlapping and non-overlapping temporal demand aggregation approaches. *Int J Prod Econ* 181:136–144. <https://doi.org/10.1016/j.ijpe.2016.04.003>
- Brady MK, Cronin JJ Jr (2001) Customer orientation: effects on customer service perceptions and outcome behaviors. *J Serv Res* 3:241–251. <https://doi.org/10.1177/109467050133005>
- Breiman L (1996) Bagging predictors. *Mach Learn* 24:123–140
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Brockwell PJ, Davis RA (2002) *Introduction to time series and forecasting*, Springer texts in statistics, 2nd edn. Springer, New York
- Brown L, Gans N, Mandelbaum A, Sakov A, Shen H, Zeltyn S, Zhao L (2005) Statistical analysis of a telephone call center. *J Am Stat Assoc* 100:36–50. <https://doi.org/10.1198/016214504000001808>
- Brown L, Cai T, Zhang R, Zhao L, Zhou H (2010) The root–unroot algorithm for density estimation as implemented via wavelet block thresholding. *Probab Theory Relat Fields* 146:401–433. <https://doi.org/10.1007/s00440-008-0194-2>
- Carboneau R, Laframboise K, Vahidov R (2008) Application of machine learning techniques for supply chain demand forecasting. *Eur J Oper Res* 184:1140–1154. <https://doi.org/10.1016/j.ejor.2006.12.004>
- Cezik MT, L'Ecuyer P (2008) Staffing multiskill call centers via linear programming and simulation. *Manag Sci* 54:310–323. <https://doi.org/10.1287/mnsc.1070.0824>
- Cleveland RB, Cleveland WS, McRae JE, Terpenning I (1990) STL: a seasonal-trend decomposition procedure based on Loess. *J Official Stat* 6:3–73
- De Livera AM, Hyndman RJ, Snyder RD (2011) Forecasting time series with complex seasonal patterns using exponential smoothing. *J Am Stat Assoc* 106:1513–1527. <https://doi.org/10.1198/jasa.2011.tm09771>
- Dean AM (2007) The impact of the customer orientation of call center employees on customers' affective commitment and loyalty. *J Serv Res* 10:161–173. <https://doi.org/10.1177/1094670507309650>
- Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc* 74:427–431. <https://doi.org/10.1080/01621459.1979.10482531>
- Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat* 29:1189–1232
- Friedman JH (2002) Stochastic gradient boosting. *Comput Stat Data Anal* 38:367–378. [https://doi.org/10.1016/s0167-9473\(01\)00065-2](https://doi.org/10.1016/s0167-9473(01)00065-2)
- Gans N, Koole G, Mandelbaum A (2003) Telephone call centers: tutorial, review, and research prospects. *M&SOM* 5:79–141. <https://doi.org/10.1287/msom.5.2.79.16071>
- Gensler S, Leeftang P, Skiera B (2012) Impact of online channel use on customer revenues and costs to serve: considering product portfolios and self-selection. *Int J Res Mark* 29:192–201. <https://doi.org/10.1016/j.ijresmar.2011.09.004>
- Holt CC (2004) Forecasting seasonals and trends by exponentially weighted moving averages. *Int J Forecast* 20:5–10
- Hyndman RJ, Athanasopoulos G (2018) *Forecasting: principles and practice*. OTexts, Melbourne
- Hyndman RJ, Koehler AB, Snyder RD, Grose S (2002) A state space framework for automatic forecasting using exponential smoothing methods. *Int J Forecast* 18:439–454. [https://doi.org/10.1016/s0169-2070\(01\)00110-8](https://doi.org/10.1016/s0169-2070(01)00110-8)
- Ibrahim R, L'Ecuyer P (2013) Forecasting call center arrivals: fixed-effects, mixed-effects, and bivariate models. *Manuf Serv Oper Manag* 15:72–85
- Ibrahim R, Ye H, L'Ecuyer P, Shen H (2016) Modeling and forecasting call center arrivals: a literature survey and a case study. *Int J Forecast* 32:865–874. <https://doi.org/10.1016/j.ijforecast.2015.11.012>
- Jalal ME, Hosseini M, Karlsson S (2016) Forecasting incoming call volumes in call centers with recurrent Neural Networks. *J Bus Res* 69:4811–4814. <https://doi.org/10.1016/j.jbusres.2016.04.035>
- Jongbloed G, Koole G (2001) Managing uncertainty in call centres using Poisson mixtures. *Appl Stoch Model Bus Ind* 17:307–318. <https://doi.org/10.1002/asmb.444>
- Kourentzes N, Petropoulos F (2016) Forecasting with multivariate temporal aggregation: the case of promotional modelling. *Int J Prod Econ* 181:145–153. <https://doi.org/10.1016/j.ijpe.2015.09.011>

- Kourentzes N, Petropoulos F, Trapero JR (2014) Improving forecasting by estimating time series structural components across multiple frequencies. *Int J Forecast* 30:291–302. <https://doi.org/10.1016/j.ijforecast.2013.09.006>
- Kourentzes N, Rostami-Tabar B, Barrow DK (2017) Demand forecasting by temporal aggregation: using optimal or multiple aggregation levels? *J Bus Res* 78:1–9. <https://doi.org/10.1016/j.jbusres.2017.04.016>
- Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R News* 2:18–22
- Mabert VA (1985) Short interval forecasting of emergency phone call (911) work loads. *J Oper Manag* 5:259–271. [https://doi.org/10.1016/0272-6963\(85\)90013-0](https://doi.org/10.1016/0272-6963(85)90013-0)
- Nikolopoulos K, Syntetos AA, Boylan JE, Petropoulos F, Assimakopoulos V (2011) An aggregate–disaggregate intermittent demand approach (ADIDA) to forecasting: an empirical proposition and analysis. *J Oper Res Soc* 62:544–554. <https://doi.org/10.1057/jors.2010.32>
- Pankratz AE (1991) Forecasting with dynamic regression models. Wiley, New York
- Parasuraman A, Zeithaml VA, Berry LL (1985) A conceptual model of service quality and its implications for future research. *J Mark* 49:41–50. <https://doi.org/10.1177/002224298504900403>
- Rausch TM, Albrecht T (2020) The impact of lead time and model selection on the accuracy of call center arrivals' forecasts. In: Proceedings of the 28th European conference on information systems (ECIS), pp 1–16
- Schapire RE, Freund Y, Bartlett P, Lee WS (1998) Boosting the margin: a new explanation for the effectiveness of voting methods. *Ann Stat* 26:1651–1686
- Shen H, Huang JZ (2008) Forecasting time series of inhomogeneous Poisson processes with application to call center workforce management. *Ann Appl Stat* 2:601–623. <https://doi.org/10.1214/08-AOAS164>
- Sitel Group (2018) 2018 CX index: brand loyalty & engagement, pp 1–27
- Taieb SB, Hyndman RJ (2014) A gradient boosting approach to the Kaggle load forecasting competition. *Int J Forecast* 30:382–394
- Taylor JW (2003) Short-term electricity demand forecasting using double seasonal exponential smoothing. *J Oper Res Soc* 54:799–805. <https://doi.org/10.1057/palgrave.jors.2601589>
- Taylor JW (2008) A comparison of univariate time series methods for forecasting intraday arrivals at a call center. *Manag Sci* 54:253–265
- Taylor JW (2012) Density forecasting of intraday call center arrivals using models based on exponential smoothing. *Manag Sci* 58:534–549
- Theodosiou M (2011) Forecasting monthly and quarterly time series using STL decomposition. *Int J Forecast* 27:1178–1195. <https://doi.org/10.1016/j.ijforecast.2010.11.002>
- Thompson HE, Tiao GC (1971) Analysis of telephone data: a case study of forecasting seasonal time series. *Bell J Econ Manag Sci* 2:515. <https://doi.org/10.2307/3003004>
- Tych W, Pedregal DJ, Young PC, Davies J (2002) An unobserved component model for multi-rate forecasting of telephone call demand: the design of a forecasting support system. *Int J Forecast* 18:673–695. [https://doi.org/10.1016/S0169-2070\(02\)00071-7](https://doi.org/10.1016/S0169-2070(02)00071-7)
- Verhoef PC, Kannan PK, Inman JJ (2015) From multi-channel retailing to omni-channel retailing. *J Retail* 91:174–181. <https://doi.org/10.1016/j.jretai.2015.02.005>
- Weinberg J, Brown LD, Stroud JR (2007) Bayesian forecasting of an inhomogeneous Poisson process with applications to call center data. *J Am Stat Assoc* 102:1185–1198
- Whitt W (1999) Dynamic staffing in a telephone call center aiming to immediately answer all calls. *Oper Res Lett* 24:205–212. [https://doi.org/10.1016/s0167-6377\(99\)00022-x](https://doi.org/10.1016/s0167-6377(99)00022-x)
- Winters PR (1960) Forecasting sales by exponentially weighted moving averages. *Manag Sci* 6:324–342. <https://doi.org/10.1287/mnsc.6.3.324>
- Young PC (1999) Recursive and en-bloc approaches to signal extraction. *J Appl Stat* 26:103–128. <https://doi.org/10.1080/02664769922692>
- Young PC, Pedregal DC, Tych W (1999) Dynamic harmonic regression. *J Forecast* 18:369–394. [https://doi.org/10.1002/\(SICI\)1099-131X\(199911\)18:6%3c369::AID-FOR748%3e3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-131X(199911)18:6%3c369::AID-FOR748%3e3.0.CO;2-K)
- Zeithaml VA, Berry LL, Parasuraman A (1996) The behavioral consequences of service quality. *J Mark* 60:31–46. <https://doi.org/10.1177/002224299606000203>

Authors and Affiliations

Theresa Maria Rausch¹  · **Tobias Albrecht²** · **Daniel Baier¹** 

Tobias Albrecht
tobias.albrecht@fim-rc.de

Daniel Baier
daniel.baier@uni-bayreuth.de

¹ University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

² FIM Research Center, University of Bayreuth and Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Bayreuth, Germany