Vaasan yliopisto
UNIVERSITY OF VAASA

Joonas Viitanen

Forecasting of inbound volumes of a call center in financial services

Case study in financial services

School of Technology and Innovation
Master's thesis in Economics and Business Administration
Industrial Management

Vaasa 2024

UNIVERSITY OF VAASA
School of Technology and Innovations
Author: Joonas Viitanen
Title of the Thesis: Forecasting of inbound volumes of a call center in financial services Case study in financial services
Degree: Master of Science in Economics and Business Administration
Programme: Industrial Management
Supervisor: Ahm Shamsuzzoha
Year: 2024 **Pages:** 65

ABSTRACT:

This study addresses the need for forecasting models in the financial sector to predict call center volumes, an essential aspect for optimizing service levels and resource allocation. The absence of a structured forecasting method in the case company's operations has created a research gap, finding out what that process should look like. With the rising significance of service level optimization, this study aimed to identify suitable forecasting models and evaluate their effectiveness, also focusing on the effect of incorporating new financing or credit application volumes into the forecasting process.

The research questions guided the investigation, emphasizing the exploration of appropriate forecasting models, determination of the best-performing model for the company, and assessment of the impact of new financing or credit application volumes on forecast accuracy. The literature review provides insights into call center operations, the importance of forecasting, and prevalent quantitative forecasting models, supplemented by prior research on workload and call volume forecasting. A quantitative methodology was adopted, utilizing data collected from the case company's SQL database tables. Based on the literarure, four time series analysis models, were chosen for comparison, tailored to the characteristics of the data and the requirement for weekly-level forecasting.

The results of the model comparison revealed that the inclusion of financing or credit application volumes did not significantly enhance forecast accuracy. Among the tested models, an autoregressive integrated moving average model emerged as the most accurate, subsequently used for constructing the forecast. This study contributes valuable insights into the application of time series forecasting models in optimizing resource allocation and enhancing service quality within call centers, particularly within the financial services industry. By addressing the identified research gap and providing practical recommendations, this research offers an addition to further advancements in call center forecasting methodologies, facilitating more efficient and effective operations within financial organizations.

KEYWORDS: financing services, time series, forecasts, contact centers, resource allocation-financing services, time series, forecasts, contact centers, resource allocation

VAASAN YLIOPISTO
Teknologian ja innovaatiojohtamisen yksikkö
Tekijä: Joonas Viitanen
Tutkielman otsikko: Puhelinpalvelukeskuksen saapuvien volyymien ennustaminen rahoitusalan toimijalle
Tutkinto: Kauppatieteiden maisteri
Koulutusohjelma: Tuotantotalous
Työn ohjaaja: Ahm Shamsuzzoha
Valmistumisvuosi: 2024 **Sivumäärä:** 65

ABSTRAKTI:

Tämän tutkielman tavoitteena oli löytää ennustemalli rahoitusalan toimijan puhelinpalvelukeskuksen volyymien ennustamiseen, sillä yritys kerää dataa palvelukeskuksensa toiminnasta mutta sitä ei tähän asti oltu käytetty tulevien volyymien ennustamiseen. Palvelutasojen optimoinnin ja resurssien kohdentamisen kasvavan merkityksen myötä tarkkojen puhelumäärien ja palvelun kysynnän ennustaminen on tullut entistä tärkeämmäksi.

Tutkimuskysymyksinä on 1. mitä ennustemalleja tulisi testata organisaatiossa, jossa puhelinpalvelun volyymien ennustamiseen ei vielä ole prosessia, 2. mikä on paras ennustemalli juuri tutkittavassa yhtiössä, sekä 3. parantaako uusien rahoitusten volyymin tai luottohakemusvolyymien huomiooon ottaminen ennusteen tarkkuutta? Kirjallisuuskatsauksessa keskityttiin puhelinpalvelukeskuksen toimintaan ja ennustamisen tärkeyteen, yleisimpiin kvantitatiivisiin ennustemalleihin, sekä aiempaan kirjallisuuteen työmäärien sekä puhelinpalvelukeskusten volyymin ennustamisesta.

Metodologia oli kvantitatiivinen ja data kerättiin yhtiön SQL tietokantatauluista ja vertailuun valikoitui neljä aikasarja-analyysimallia niiden sopiessa parhaiten käytettyyn dataan sekä viikkotason ennustamiseen. Aikasarjamallien vertailun tulosten perusteella rahoitus- tai luottohakemusvolyymien vaikutus ei ollut merkitsevä ja vertailun tarkin oli ARIMA malli, jolla ennuste rakennettiin. Yhteenvetona tämä tutkimus tarjoaa näkökulmia aikasarjaennustemallien soveltamiseen resurssien kohdentamisen optimoimiseksi ja palvelun laadun parantamiseksi puhelinpalvelukeskuksessa, erityisesti rahoitusalalla.

AVAINSANAT: rahoituspalvelut, aikasarjamallit, ennustaminen, puhelinpalvelukeskus, resurssien allokointi

Contents

Figures

Tables

1 Introduction

“We should all be concerned about the future because we will have to spend the rest of our lives there.”
- C.F. Kettering (Andrews, 1989, page 105).

As the world becomes more and more reliant on data, it is increasingly important to utilize the collected data to improve a company’s operations, and call centers are definitely not an exception. One essential use for this data is the ability to predict future activity, to be able to maximize efficiency in resource planning and increase customer satisfaction. In fact, being able to accurately forecast call center activity is one of the most important steps in improving a call center’s operations (Aksin et al., 2007; Cleveland, 2019; Gans et al., 2003; Saccani, 2013) and has been for a long time (Bodin, 1999).

As a result, a comprehensive body of knowledge has accumulated around different forecasting methods used in call centers (Antipov & Meade, 2002; Barrow, 2016; Bugarčić & Radojičić, 2020; Ding & Koole, 2022; Gans et al., 2015; Ibrahim et al., 2016; Rausch et al., 2022; Saccani, 2013) and for forecasting time series data in general (Bertsimas & Boussioux, 2023; Box et al., 2015; Cryer & Chan, 2010; Hyndman et al., 2011; Hyndman & Athanasopoulos, 2021; Makridakis et al., 1997; Oliveira et al., 2014). The benefits of good call center forecasting are numerous. In addition to improved resource allocation, staffing, and customer experience (Aksin et al., 2007; Mehrotra & Fama, 2003), a volume forecast also provides the company and call center management with tools for service level optimization and possibly better cost management, as they are able to avoid over and understaffing which often lead to reactionary decisions and therefore to extra costs (Aksin et al., 2007; Cleveland, 2019).

1.1 Aim of the thesis

The aim of this thesis is to build a predictive model for the call center of a financial services company using predictive analytics methods. Using this model, the organization would be able to transition from a reactive approach to operational resource allocation to a more proactive one in handling inbound calls. The research is conducted as a quantitative study.

In the current situation the company collects data on call center activities, but the data is utilized in visualizing historical and current workloads, with no method for forecasting into the future. However, due to the availability of the data on both the calls received as well as some factors that the company hypothesizes might be driving changing levels in the call center volumes, it is possible to create a model to forecast this volume. These factors are the number of financed agreements, and number of inbound credit applications, as both are indicative of new business and therefore potentially new customer contacts.

This results in a fillable research gap where necessary data exists and is collected but is not used for predictive purposes due to lack of a process. This thesis aims to formulate a prediction process and allowing the organization to move from a reactive operations resource management strategy to a proactive one. The main objective therefore is to be able to present company management with a decision support tool for operational resource planning.

The model is built using predictive analytics, which is a broad category of data analysis techniques ranging from time series analysis to machine learning, artificial intelligence, and pattern recognition, in which the model itself determines the key factors, such as variable weights and coefficients, as opposed to the model user's assumptions and inputs (Abbott, 2014).

1.2 Research questions

The research questions this thesis aims to answer are as follows.

1. What are the models that should be tested in an organization with no existing forecasting capabilities in their call center operations?

2. What is the best model for building a call center volume forecast in the case company specifically?

3. Do new financed agreement volumes or inbound credit application volumes predict call center volumes sufficiently to be used in the forecast?

1.3 Thesis structure

This thesis is divided into five parts. First, the introduction chapters lay out the subject of the thesis, the research questions, as well as give an overview of what the thesis contains.

Second, a literature review is conducted, diving into the working of call centers in general, call center forecasting principles as well as predictive analytics methods including time series analysis, neural networks, linear regression, and ensemble methods. The literature into call centers especially highlights how the problems call centers face are in part still the same as in the past (Bodin, 1999; Cleveland, 2019), namely, how to forecast and predict volumes for efficient resource management and planning. Different types of call center inbound volumes and the factors that influence them are also presented as well as their importance to companies' value creation.

In terms of predictive analytics and time series analysis, different methods and best practices are reviewed (Abbott, 2014; Box et al., 2015; Hyndman & Athanasopoulos,

2021) and various practical examples are presented from different fields where these methods have been used to make forecasts in the call center context (Antipov & Meade, 2002; Barrow, 2016; Bugarčić & Radojičić, 2020; Ding & Koole, 2022; Gans et al., 2015; Rausch et al., 2022; Saccani, 2013) as well as other fields (Anglou et al., 2021; Edna Chan et al., 2002; Fijorek & Leśniewska, 2013; Piccialli et al., 2021; Teoh Yi Zhe & Keikhosrokiani, 2021; van Gent et al., 2018).

Third, the methodology of the thesis is presented, including how the data used has been selected and prepared, what tools have been used in the analysis and what forecasting methods are compared to find the right model. This section describes how the case company collects call center related data into databases, how far back the data is available, which external variables are analyzed alongside the call center volume and how the data is aggregated to be able to create a forecast of the desired granularity. The analysis tool, Alteryx software, and the reasons behind its use are also presented.

In the fourth part the results of the data analysis are presented, divided into an exploratory analysis phase and the model comparison phase. The fifth and final part contains the conclusions that can be drawn from the results as well as a discussion on the limitations of the thesis as well as potential further research opportunities. This is followed by a list of references used, as well as an appendix with additional graphs.

1.4 Case company

The case company is an international financial services provider. The company has 5000+ employees and over 5 billion euros in revenue with a strong presence in the Nordics. It provides a comprehensive range of financial services, including retail banking, corporate banking, investment banking, asset financing, and wealth management. The organization serves various customer categories, ranging from individuals and small businesses to large corporations and institutional clients.

This study focuses on the call center of one specific product area inside the company. The product area in question, asset finance, deals with leasing and hire purchase products for both corporate and private customers in Finland.

Due to the nature of the financial services industry, it has been requested by the case company that specific call, agreement, and application volume figures are not disclosed in this thesis, but instead their values normalized on a 0-1 scale where such figures are presented, 1 being the highest value of the data and 0 the lowest.

2 Literature review

The literature review consists of three main topics. First, an introduction into the functioning of call centers and the importance of forecasting in their improvement and management. Second, a review of predictive analytics and forecasting methods and best principles. Finally, existing studies that focus on workload forecasting and then more specifically call center volumes. The implications for this thesis according to the findings from the literature review are also presented.

2.1 Call center operations

Customer service is an essential part of almost every company's operations and even more so for companies providing services in a more and more complex and digital environment. A dedicated call center is therefore a necessity for most companies serving a large number of customers and companies working in the financial services industry are definitely not an exception.

According to customer service expert Brad Cleveland call centers are still struggling with the same problems now as when the first ones were established in the beginning of the 20th century (Cleveland, 2019). Namely, the problem or resourcing, or how to know when demand for call center personnel is at its highest to be able to meet demand without over-resourcing when demand is lower. And while technology and self-service channels have improved to a point where easy to solve issues can be handled in some cases by the customer themselves, there is no less need for great call center processes, as the issues handled by agents are now often more complicated and the impact on a company's reputation and customer satisfaction is as high as ever.

A focus on running a well-oiled and efficient call center can even bring strategic value to an organization by increasing loyalty through customer satisfaction. This is especially important in the financial services industry, where products are often close to identical due to regulations, so brand image and customer experience play a significant part in attracting and retaining customers.

Figure 1. Strategic value through call center efficiency (Cleveland, 2019)**.**

According to the International Customer Management Institute, call center management is even defined as “the art of having the right number of properly skilled people and supporting resources in place at the right times to handle an accurately forecasted workload, at service level and with quality.” (Cleveland, 2019, p. 22). Therefore, it can be derived that forecasting call center workloads is one of the most essential functions in customer center operations. While in previous decades getting access to historical call center data was a sought-after way of maximizing customer service efficiency and finding patterns in caller behavior (Bodin, 1999), with organizations’ current data capabilities forecasting is the logical next step.

Cleveland defines three main driving forces in determining call center volumes and management practices and stresses that understanding these three forces is key to making scheduling and staffing improvements (Cleveland, 2019, p. 50).

1. Workload arrival
2. The queue
3. Customer tolerance

From a forecasting perspective the most significant is the first one, workload arrival. Cleveland categorizes arrivals into three types: random, smooth, and peaked. Effectively these categories try to explain the trend in customer contacts, random meaning essentially that the one customer contact is not correlated with the likelihood of the next one. Smooth on the other hand exhibits a clear trend over time but is usually more associated with outbound calls according to Cleveland, as opposed to inbound calls. Finally, peaked arrival describes a situation where the contacts surge around a specific peak time, whether that be in a certain weekday, hour inside a day, or other short period of time.

Figure 2. Three types of workload arrival (Cleveland, 2019)**.**

When it comes to forecasting call center contacts with quantitative methods, Cleveland identifies two different routes. Either explanatory forecasting, for example using regression or multivariate methods, or time-series analysis which is found sufficiently accurate for shorter term forecasts of up to three months forward. Explanatory methods are favored when one is looking for correlations to external factors that might influence call center volumes, such as a price increase.

2.2 Predictive analytics and forecasting

Predictive analytics is characterized by two things separating it from other types of analytics. It is based on data, in other words the user is relying on the characteristics of the data itself instead of their own assumptions in creating the model, and it is forward looking, as can be inferred from the name. While models differ greatly in how they work internally, they all more or less follow an assumption that what happened in the past will continue to happen in the future (Abbott, 2014).

However, this does not mean that forecasts cannot be used in a changing environment, and a correction could be made that the forecaster assumes that the environment will keep changing the same way (Hyndman & Athanasopoulos, 2021). The process of building and maintaining a predictive model also often follows the same timeline, pictured below, where the model is calibrated continuously after new data is available for use within the analysis at a later date.

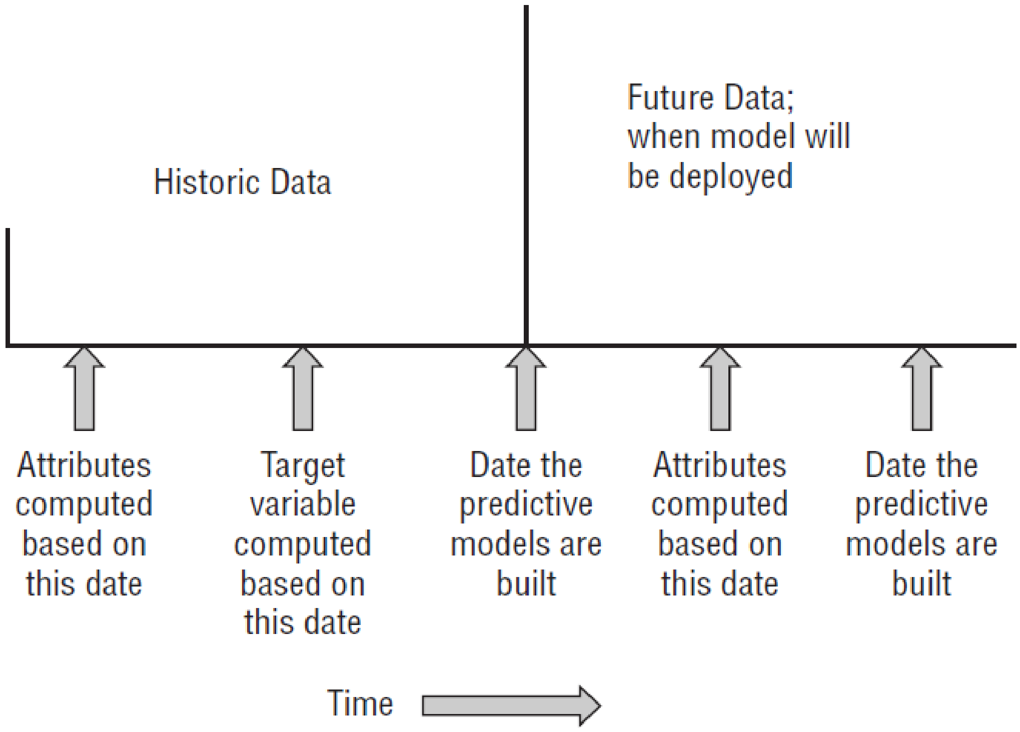

Figure 3. Timeline for building predictive models (Abbott, 2014)**.**

According to Rob Hyndman and George Athanasopoulos in the introduction to their book Forecasting: Principles and Practice (2018), humankind has sought to predict the future in one form or another since ancient Babylon. He defines forecasting as “predicting the future as accurately as possible, given all of the information available, including historical data and knowledge of any future events that might impact the forecasts.” (R. Hyndman & Athanasopoulos, 2021, p. 4). When deciding to create a quantitative forecast, numerical data of the past must be collected and available, and there must be an assumptions that patterns found in the data will continue in the future.

If these conditions are satisfied, a preliminary data analysis should be conducted, where the data is graphed, and some trends or features might be visible to the eye and help in choosing the models to test. The next step is deciding which models to try in finding the best possible forecast.

Usually at least two or three different models should be tested (Hyndman & Athanasopoulos, 2021) and using diagnostics and error terms the best fitting one can be determined. This also applies to the parameters inside models. The ARIMA time series model, for example, should be tested with different degrees of differencing and orders of the moving average and autoregression components (Box et al., 2015).

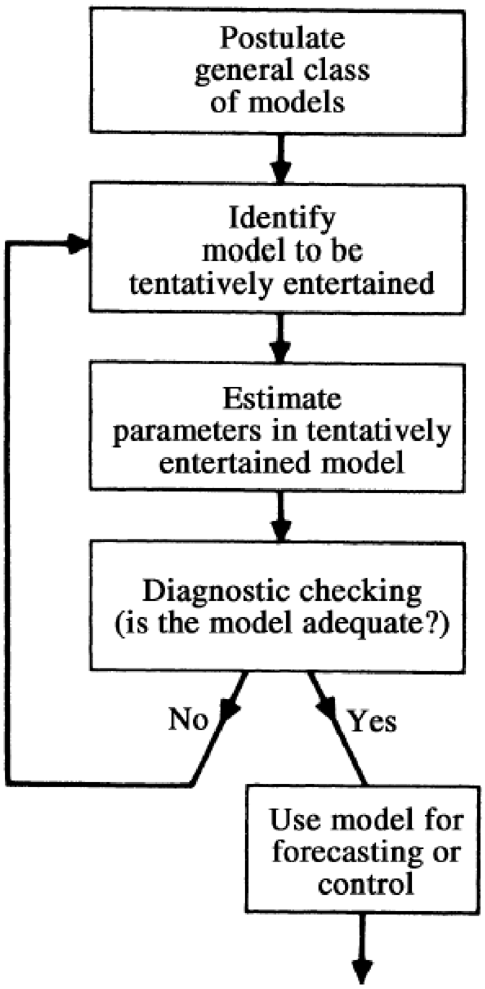

Figure 4. Stages in the iterative approach to model building (Box et al., 2015)**.**

One of the most severe challenges in building a forecasting model is avoiding overfitting the data by building a model that is too sophisticated right out of the gate, since the error of overfitting is most likely only noticed after the model has already been deployed and in use (Abbott, 2014). Overfitting means the model includes more complexity than is actually needed, either by including too many terms compared to the need or too much flexibility, for example by using a neural net to model a linear relationship that would be better served by a regression model. These issues can lead to inaccurate predictions, as irrelevant predictors create unneeded variation and makes the model less verifiable by others due to the complexity (Hawkins, 2004).

2.3 Forecasting techniques

In this chapter we will go through the most common analytics techniques identified by the Rexer Analytics data science survey in 2020, a comprehensive survey on predictive analytics (Abbott, 2014). 579 data science professionals were surveyed from 71 countries on, among other things, which algorithms they use in their fields, both industry and academia. The ten most common answers are presented in figure 5. The algorithms that are not forecasting methods, text mining, cluster analysis and principal component analysis, are however not covered in this thesis.

Decision trees on the other hand are partitioning algorithms, essentially a series of *if else* statements that are meant to identify a value that is being predicted, whether that is a categorization or prediction for example (Kuhn & Johnson, 2016). The popularity of decision trees mostly stems from them being easy to build and understand (Abbott, 2014) but while decision trees can be used for predicting values, they are not ideal for handling time series data, although they have been used to choose the best time series model with encouraging results (Jakobs & Saadallah, 2023). Logistic regression is also primarily a classification model and as such, not ideal for forecasting from time series data (Abbott, 2014).

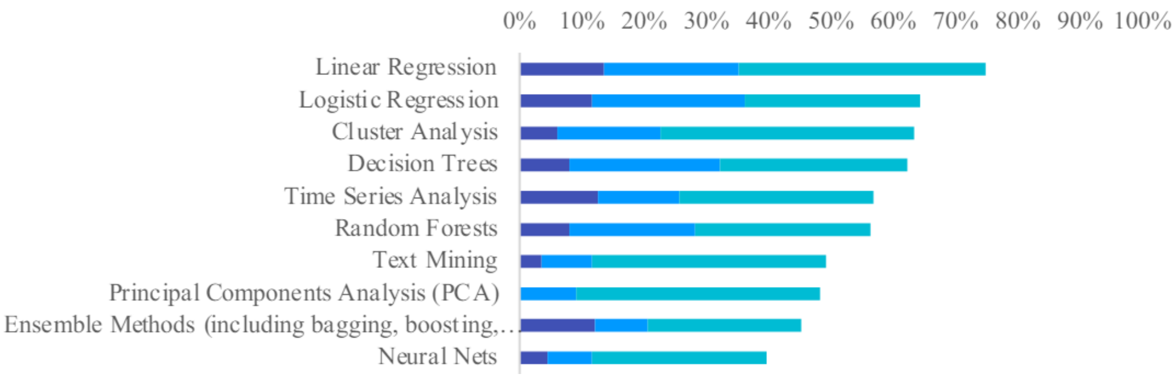

Figure 5. Most commonly used data science algorithms (Rexer, 2020)**.**

2.3.1 Linear regression

According to Rexer's survey, linear regression is the single most widely used data science algorithm. Linear regression is a method of supervised learning that aims to find parameters with which the sum of squared errors is as small as possible in finding a relationship between variables (James et al., 2013).

The linear regression equation is described as follows (Kuhn & Johnson, 2016):

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdot\cdot\cdot + b_P x_{iP} + e_i, \quad (1)$$

where

y is the dependent variable

b0 is the estimated intercept

bj is the estimated slope coefficient for the *jth predictor*

xij is the predictor variable for the *j*th predictor and *i*th sample

e is the error term

When using linear regression there are some assumptions that need to be made for it to be a viable method for analysis (Abbott, 2014). A linear relationship between the target and input variables is assumed, and the inputs should not be correlated with each other. The model residuals are also assumed to be normally distributed and having a mean of zero.

2.3.2 Time series analysis

When referring to time series analysis, what is most often meant is various smoothing methods that are based on a moving average (Shmueli & Lichtendahl, 2016). The most common of these are the ARIMA and exponential smoothing models, being covered in most time-series analysis textbooks (Box et al., 2015; Cryer & Chan, 2010; Shmueli & Lichtendahl, 2016).

Exponential smoothing (or ETS) is a widely used time-series forecasting technique that is based on the idea of giving more weight to recent observations than to past ones and is a recursive formula that uses a weighted average of past observations to forecast future values of the time series. The weights assigned to each observation decrease exponentially as the observations get further into the past. The model can be formulated as

$$Y\_t+1 = \alpha * Y\_t + (1 - \alpha) * F\_t. \tag{2}$$

Y_t+1 is the forecast for the next period, while α is the smoothing parameter, Y_t is the actual observation in the current period, and F_t is the forecast for the current period. The smoothing parameter α determines the weight given to the most recent observation, and it is typically chosen through cross-validation or optimization techniques.

Exponential smoothing models can be further classified based on the number of smoothing parameters used. Single exponential smoothing (SES) uses one smoothing parameter and is suitable for time series with no trend or seasonality. Double exponential smoothing (DES), also known as the Holt-Winters model, uses two smoothing parameters and can capture both trend and seasonality in the time series. Triple exponential smoothing (TES), also known as the Holt-Winters-Additive model, uses three smoothing parameters and can handle time series with trend, seasonality, and a changing level of volatility (Hyndman et al., 2011).

Exponential smoothing has been shown to perform well in a wide range of applications, such as forecasting sales and stock prices (Hyndman et al., 2002; Makridakis et al., 1997). The method is particularly useful when the time series exhibits trend, seasonality, or a combination of both, as it can capture the underlying patterns in the data and provide accurate forecasts. In conclusion, exponential smoothing is a powerful and widely used time-series forecasting technique that can capture the underlying patterns in the data and provide accurate forecasts.

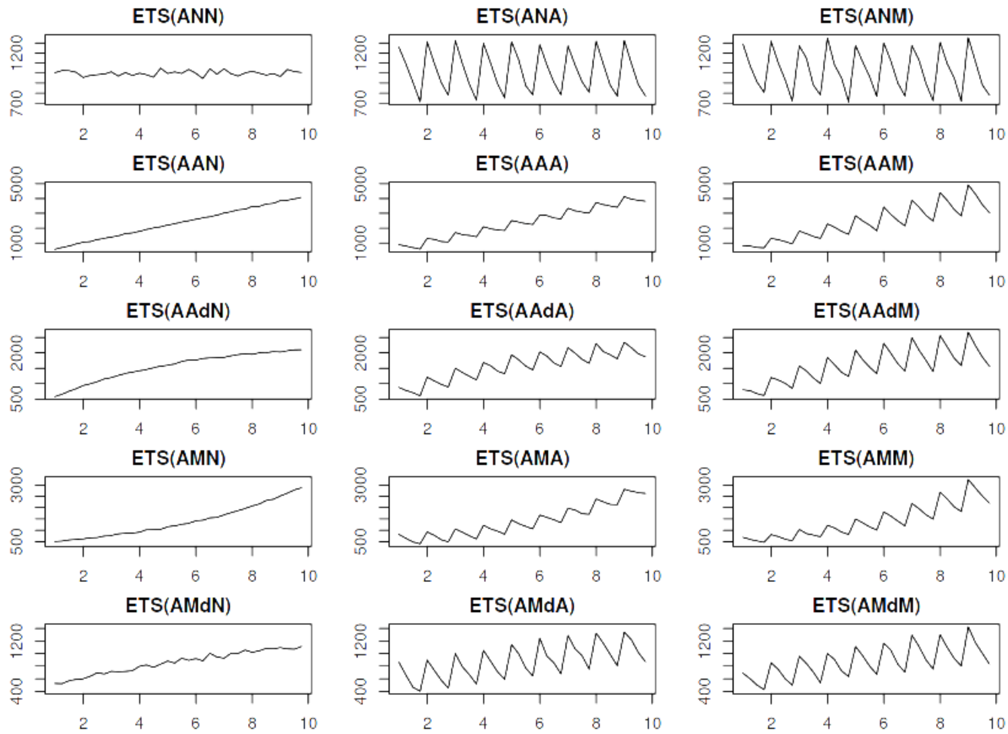

Figure 6. Time series corresponding to the additive error ETS models (Svetunkov, 2023)**.**

Svetunkov in his 2023 book presents a convenient table for a graphical rule of thumb categorization on what type of model generally fits a time series dataset with certain characteristics. This can be seen in figure 6. For example, a stationary data series with no trend or seasonality, as in the first graph, is best fitted by an ETS(ANN) model. The three letters correspond to the types of error, trend, and seasonality, where ANN would mean that error is additive, but there is no trend and no seasonality, while an ANM model for example represents a model where the error is again additive and there is no trend, but there exists a multiplicative seasonal component (Hyndman et al., 2008). In other words, ETS = **E**rror **T**rend **S**easonality, where error can be either A = additive or M = multiplicative, trend can be N = none, A, M, Ad = additive damped, or Md = multiplicative damped, and seasonality can be N, A, or M. An additive term means that the values are assumed to stay the same throughout the series, while multiplicative terms are assumed to vary according to the patterns in the data.

ARIMA (Autoregressive Integrated Moving Average) models are another popular time-series forecasting method used in various fields, including finance, economics, and engineering.

The ARIMA models a time series as a combination of its past values and random error terms. The model consists of three components: autoregression (AR), differencing (I), and moving average (MA). The AR component models the dependence between the current observation and its past values, the I component is used to make the time series stationary, and the MA component models the error term.

The order of the ARIMA model is specified as (p, d, q), where p is the order of the AR component, d is the degree of differencing, and q is the order of the MA component. ARIMA models have been shown to perform well in a wide range of applications, such as forecasting stock prices and exchange rates (Bollerslev et al., 1992; Makridakis et al., 1982). It is also possible to consider covariates outside the past forecast value data points (Box et al., 2015) in what is often called ARIMAX (ARIMA with exogenous variables).

The model is defined as follows:

$$Y_t = c + \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + \theta_1 e_{t-1} + ... + \theta_q e_{t-q} + X_t \beta + e_t, \quad (3)$$

where

Y_t is the value of the time series at time t.
c is a constant term.
$\phi_1, ..., \phi_p$ are the autoregressive coefficients.
$Y_{t-1}, ..., Y_{t-p}$ are the lagged values of the time series.
$\theta_1, ..., \theta_q$ are the moving average coefficients.
$e_{t-1}, ..., e_{t-q}$ are the lagged errors.
X_t represents the exogenous variables at time t.
β represents the coefficients of the exogenous variables.
e_t is the error term.

In conclusion, ARIMA and ARIMAX models are popular and effective time-series forecasting techniques that have been widely used in various fields. ARIMAX models are particularly useful when external factors need to be considered in the forecasting process.

2.3.3 Ensemble methods

Model ensembles are when a collection of multiple forecasting techniques are used in building a single forecast. The reason for doing this is that several models working in tandem might, and often does, bring about better results than relying on a single model, as listed in figure 7 (Abbott, 2014). Sacrifices have to be made however as some of the interpretability is lost as the model grows more complex and becomes more expensive to run in terms of computing power, making the forecasting process slower than a single model equivalent.

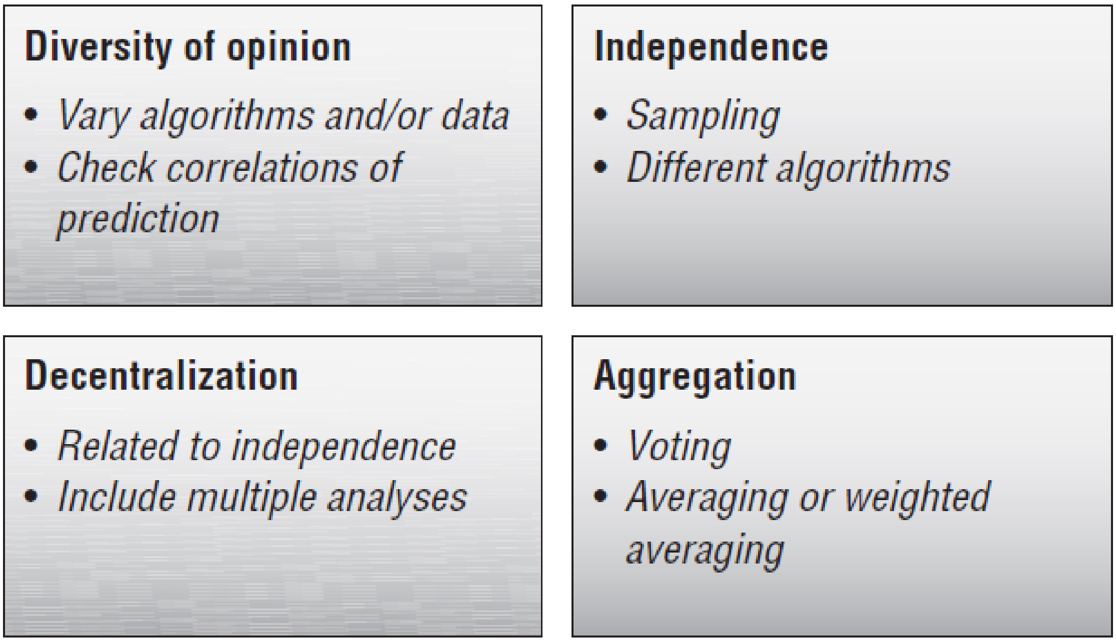

Figure 7. Characteristics of good ensemble decisions (Abbott, 2014)**.**

Abbott divides model ensembles into bagging and boosting. Bagging, or bootstrap aggregating, was initially developed for decision trees but is applicable to many other forecasting models also. It works by building multiple forecasts from bootstrapped samples and averaging the results to get a more accurate representation of the data. Boosting on the other hand tries to improve poor forecasts by combining them and creating a more robust learning model from several weaker ones. Boosting is mostly used with decision trees and in classification problems (Abbott, 2014).

Ensemble methods have also been studied in time series data forecasting, trying to improve time series models in choosing the best among multiple models over longer periods of time, as the best model at one point in time might not be the best model also in the future if trends or other characteristics of the data change (Oliveira et al., 2014). While the results were encouraging, the temporal nature of the data does add complexity to using ensembles, leading to further researching in building more robust ensembles using adaptive robust optimization, again improving forecasts but requiring a large training set of at least several hundred data points (Bertsimas & Boussioux, 2023).

2.3.4 Neural nets

Artificial intelligence algorithms, and specifically neural nets, are another popular forecasting tool, which is also relatively widely used for forecasting time series (Shmueli & Lichtendahl, 2016). While at its core neural networks are based on the linear regression model, the way in which a neural network functions allows it to be used in modeling non-linear relationships.

This is achieved by emulating the human brain, consisting of interconnected neurons, or nodes, organized into layers. Data flows from the input layer through hidden layers to the output layer, with each neuron processing inputs and applying an activation function to introduce non-linearity (Shmueli & Lichtendahl, 2016). The network learns through a process called backpropagation, where it adjusts the weights of connections based on the error between its predictions and actual target values, as measured by a loss function (Goodfellow et al., 2016).

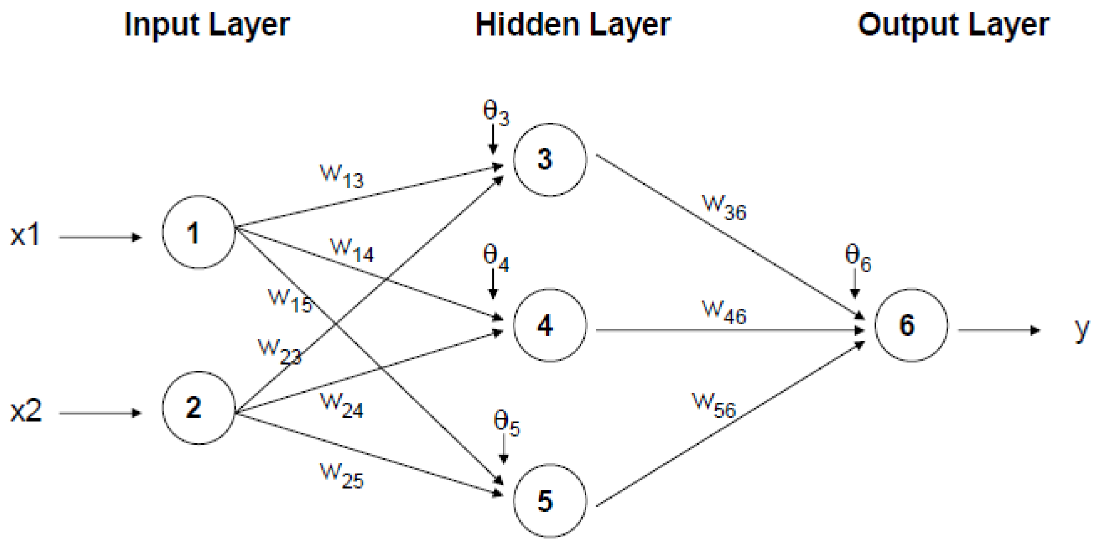

Figure 8. Schematic of a neural network with a single hidden layer, for predicting a single outcome (y) from two predictors (x1, x2) (Shmueli & Lichtendahl, 2016).

The benefits of neural networks compared to other models is when capturing complex relationships between variables. Better results in time series forecasting have been acquired when an abundance of data points is available, for example in intra-day forecasts and when used in an ensemble model (Shmueli & Lichtendahl, 2016). In other use cases results have been inconclusive.

2.4 Examples of workload forecasting

Very little academic literature seems to exist on financial services operations forecasting (Hatzakis et al., 2010), most likely due to the differences in characteristics of various processes. The operational workload associated with financing a car for a private customer, for example, is determined by the sophistication of the customer portal (end-user portal versus manual work from customer service personnel), amount of manual work in the financing process etc.

Therefore, as the processes, systems, and levels of automation are not standard across the industry, workload forecasting is always situation dependent. It is possible despite this to draw from studies in other industries, where operations and workload forecasting is more prevalent.

Some studies predicting the mental workload of knowledge workers have been conducted, making use of both self-assessment and physiological trackers, the data being analyzed using machine learning (van Gent et al., 2018) and deep learning (Teoh Yi Zhe & Keikhosrokiani, 2021), which are more suitable for data with high dimensionality and an abundance of data points.

While machine learning and deep learning models are generally the superior choice when dealing with “big data”, the nature of the data in this study is such that even with multiple dimensions the number of data points does not necessarily rise to a sufficient level. Additionally, this study is more focused on predicting the workload of the whole team, instead of individuals.

Such a team-centric study was conducted in a hospital medical records department, by Chan et. Al in 2002. By quantifying the time it takes to perform various workflow steps, the authors were able to model optimal staffing rotations as well as provide a modeling tool for the management to support in their resource optimization decisions using ARIMA, or Autoregressive integrated moving average. While the workloads experienced high variability from week-to-week and month-to-month, the forecast simulation was useful in seeing the average effect of changes in resource management decisions (Edna Chan et al., 2002).

A similar study conducted in an Italian health department tested machine learning methods, specifically boosted decision trees, random forest algorithms, and recurrent neural networks, in addition to ARIMA, in this case incorporating seasonal and exogenous variables (Piccialli et al., 2021). The authors’ final forecast used a combination of methods for a new ensemble technique tailored for workload prediction in a setting where abundant and detailed data is available for use.

A field in which workload forecasting using analytics has been especially prevalent is in supply chain management. Some use cases include predicting spare-part needs in the shipping industry (Anglou et al., 2021) and optimal resource allocation in warehouse order-picking (van Gils et al., 2017).

The authors in the 2021 study used clustering techniques combined with machine learning in their predictive model to first identify the most high-impact parts for ordering via clustering, and then compared different machine learning techniques to assess which is the most suitable for predicting when parts are needed and in what quantities. From the tested methods of Random Forest, Generalized Linear Model, and Principal Component Regression, the random forest method was found to be the most optimal, due to its handling of stochastic behavior.

Van Gils et al. on the other hand relied on multiple types of time series forecasting in predicting daily order picking volumes in a warehouse, in order to make staffing decisions on a daily basis. A mix of SARIMA, exponential smoothing, and combination forecasts proved most accurate in the study, due to considering trends and seasonality, which are most likely also a feature in the data of this study. Similar methodology has been employed in studying airport operations forecasting (Fijorek & Leśniewska, 2013).

2.4.1 Call center demand forecasting

The most relevant studies in this paper's context, however, are those conducted on call center demand forecasting. These seem mostly to follow the most popular time series forecasting methods to determine the best model for forecasting.

The field of call center forecasting can be considered quite new, since until the 2000's the amount of research was very limited, one possible reason for this being that while call centers in theory generate several data points for each call, such as the caller info, call length etc., these have not always been recorded to a database due to cost of storage and system capabilities (Gans et al., 2003). Forecasting has been identified as one of the most important operational issues in call center management in literature reviews of the call center management field (Aksin et al., 2007), with a role still growing in importance in the future.

Adopting a rigorous process for creating demand forecasts can be a major first step in call center capacity and planning improvements, especially in an operation without any previous forecasting process in place. Saccani (2013) divides call center capacity planning into shorter term forecasts that are concerned with capacity planning in the next few hours, or even shorter time frames, and he calls these queuing methods.

The more interesting for this thesis is what Saccani calls scheduling methods, which are forecasting ahead for days or weeks and which can be used, for example, in planning recruitments. The data also looks very different based on whether we are forecasting on the weekly or hourly timescale for example, as can be seen in figure 9.

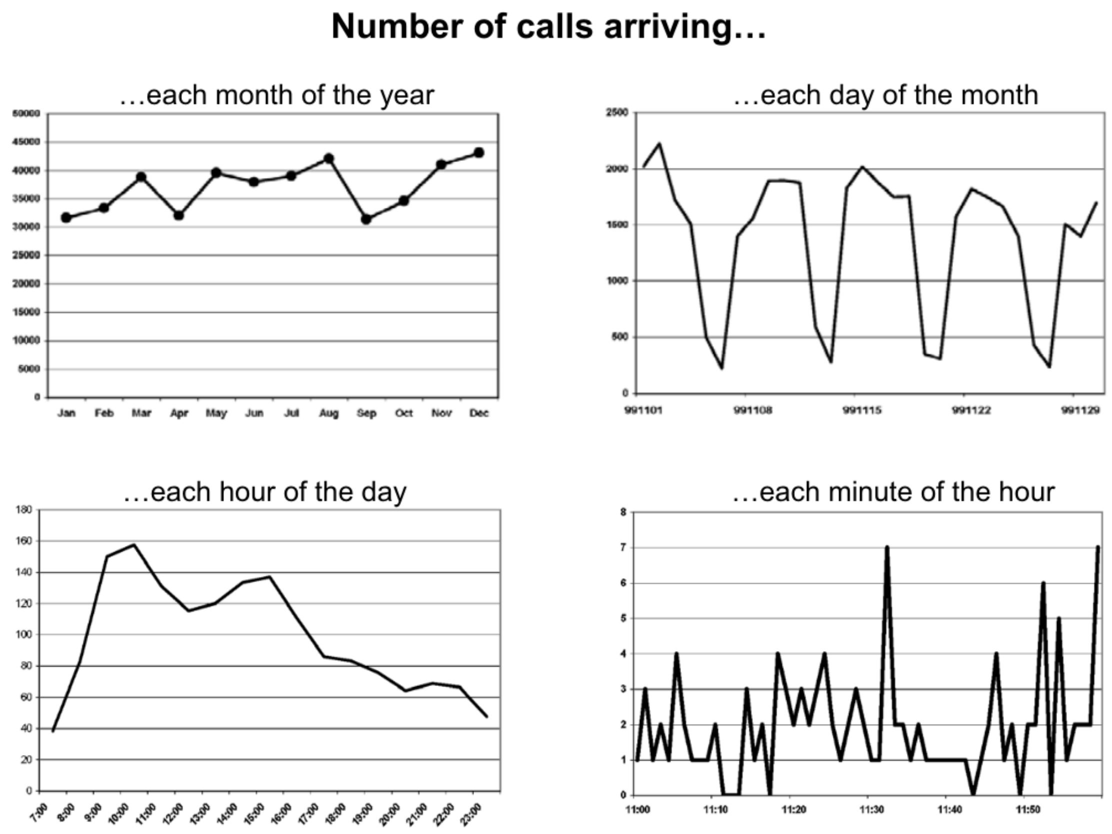

Figure 9. A Hierarchical View of Arrival Rates (Gans et al., 2003)**.**

The implementation of the model in the organization can also be more crucial than the sophistication of the model itself, since it has been found that simple forecasting techniques can perform as well as those that are more complex (Saccani, 2013), especially for longer forecasting periods (Ibrahim et al., 2016). This is supported by some of the actual results of call center demand forecasting studies, where rather common time series forecasting methods such as Holt-Winters (Barrow, 2016; Bugarčić & Radojičić, 2020), and ARIMA (Antipov & Meade, 2002; Barrow, 2016) models were the best performing out of those chosen by the authors, especially in the short term. Some research has been done into more complex models such as stochastic programming and recourse (Gans et al., 2015) or models that are call center specific such as Erlang C (Ding & Koole, 2022). However, the majority seem to favor the more common suite of time series forecasting methods (Rausch et al., 2022).

Based on a literature survey on call center analysis methods used to create forecasts (Ibrahim et al., 2016), including some studies mentioned in this thesis, the ARIMA and exponential smoothing methods were the most favored by researchers. A summary of the literature survey result, amended with additional studies identified by Barrow in 2016 and found by the author of this thesis, are presented in table 1.

Table 1. Review of call center forecast models used in previous research

Authors	Year	Time-series	Regression	Bayesian	Linear fixed effects	Artificial neural networks	Other, what?
Aktekin and Soyer	2011			x			
Aldor-Noiman et al.	2009				x		
Andrews and Cunningham	1995	x					
Antipov and Meade	2002	x					
Avramidis et al.	2004						Dirichlet distribution
Barrow	2016	x				x	
Bianchi et al.	1998	x					
Bugarčić and Radojičić	2020	x					
Channouf	2007	x					
Channouf	2008						Copula model
Ding and Koole	2015						LP optimization
Ding and Koole	2022						Erlang C
Gans et al.	2015						Stochastic programming
Ibrahim and L'Ecuyer	2013				x		
Jongbloed and Koole	2001						Dirichlet distribution
Mabert	1985	x	x				
Oreshkin et al.	2016						Copula model
Shen and Huang	2008				x		
Soyer and Tarimcilar	2008			x			
Taylor	2008	x			x		
Taylor	2003	x					
Taylor	2010		x				
Taylor	2012	x					
Thompson and Tiao	1971	x					
Kim et al.	2012				x		Copula model
Zhang	2003	x				x	
Tych et al.	2002		x				
Taylor and Snyder	2012					x	
Pacheco et al.	2009					x	
Weinberg et al.	2007			x	x		
In total		**12**	**3**	**3**	**6**	**4**	**8**

The “time-series” column refers to ARIMA and exponential smoothing methods. In the literature survey, another very popular forecasting method, developed by multiple authors in succession, was the linear fixed effects method, which aims at modeling the arrival rates of calls based on a Poisson distribution. These are especially popular for interday and intraday forecasts, while for longer time periods time-series methods were favored (Ibrahim et al., 2016).

In general, it seems that attention towards more complex models in call center forecasting have been reserved for daily and intra-daily forecasts, instead of more aggregated weekly or monthly forecasts (Saccani, 2013). Another popular forecasting method has been utilizing artificial neural networks, which has shown mixed results according to the literature reviewed by Barrow (Barrow, 2016), and is also focused primarily on intraday forecasts.

2.5 Implications for the thesis

Based on the existing literature, it seems clear that being able to produce even a rudimentary forecast for customer service operations is one of the most important steps in improving efficiency (Cleveland, 2019; Saccani, 2013), and has been identified as such for a long time (Bodin, 1999). As the case company already has data tables available that records call center activities, it seems like an obvious progression to utilize the data not only for historical tracking, but for looking forward as well.

While workload forecasting in general and call center volume forecasting in specifically has a wide array of predictive analytics methods to choose from based on existing studies, favoring simplicity in the initial forecast seems to bring close to the same amount of benefits as a more complex model, particularly when aggregating the data to a weekly forecast (Ibrahim et al., 2016; Saccani, 2013). For this reason and due to the nature of the data (limited number of data points), common timeseries analysis is chosen as the preferred method in this study, supported by results of several similar studies in other industries dealing with call center volumes (Antipov & Meade, 2002; Barrow, 2016; Bugarčić & Radojičić, 2020) as well as them being identified as often superior to more complex models when forecasting on a longer time horizon than daily or intra-day forecasts.

While ensemble methods and neural networks have also seen encouraging results in time series forecasting (Bertsimas & Boussioux, 2023; Ibrahim et al., 2016; Oliveira et al., 2014; Shmueli & Lichtendahl, 2016), a simpler model to start with is preferred, both for the aforementioned reasons as well as to ensure that the model can actually be maintained and adjusted relatively easily in the organization, as forecasting is often and also in this case an ongoing process instead of a single event (Shmueli & Lichtendahl, 2016), and to avoid overfitting by adding complexity for the sake of complexity.

3 Methodology

The methodology section is divided into three parts and includes an introduction to the data used in the analysis, the tools used for data manipulation and forecasting, as well as the decision on what forecasting models are being compared. This study uses a quantitative methodology, as it provides a structured, objective, and rigorous framework essential for developing, testing, and validating forecasting models. Quantitative methods’ reliance on numerical data and statistical techniques ensures accuracy, reliability, and the ability to generalize findings.

3.1 Data selection

The case company collects data from their contact center services platform into a relational SQL (structured query language) database. This call center data, a normalized sample of which can be found in the appendix, is used as the basis of the study. Additionally, the call center volume data is tested alongside two different external variables that the author hypothesizes could have an effect on the changing volumes. These are incoming credit applications and started financing agreements, as it is worth investigating whether new agreement volumes also drive customer service activity. The data for the above variables can be found in another SQL database connected to the company’s Enterprise Resource Planning system which records agreement and credit application data.

Due to organizational data retention rules, the dataset history is limited to 166 weeks of data, of which the full scope is used to train and test the models. The data used starts on 18.1.2021 and ends with the week starting 18.3.2024. With special permission, it would be possible to attain a longer period of data, but only for a point-in-time analysis, instead of a continuous process where the forecast is updated weekly.

For that reason, the data used will only cover what is within data retention rules in the company’s structured data mart. To be able to have a sufficient number of observations, the forecasted time period should be either weeks or days, as months would provide less than 50 observations.

The needs of the forecast model from the organization are focused on allocating resources for the next weeks as well as to allow for a ramp-up of personnel ahead of time, so daily or hourly forecasts are not crucial. For these reasons, the data for model development constitutes just over three years (166 weeks) of historical data and observation frequency is aggregated to a weekly level to be able to provide weekly forecasts.

3.2 Data preparation and analysis tools

The analysis tool used both for creating the forecast as well as data wrangling is Alteryx Designer, an analysis software application for Windows, used as the main analysis tool in the case company. There are two main reasons this tool is selected. As the forecast creation process is not designed to be a one-off, but a continuous weekly forecast, the tool to run it should be one that has a wider user base in the company and that is approved by the company’s IT department for use in business processes.

The second reason is that the tool has an intuitive user interface and is able to handle both the data preparation and running of different predictive models with very little coding knowledge needed, which makes the forecast easier to maintain in the future. The Alteryx Designer time series tools are built on top of the programming language R, a language specifically designed for statistical computing (Cryer & Chan, 2010) often used in time series and other statistical analyses.

The same logic applies for the choice of forecast presentation, which is done with the Tableau business intelligence and analytics software. As Tableau is the primary business intelligence tool in the case company, the forecast report published to the contact center personnel is built with Tableau.

The call center data warehouse includes data from all group companies and different departments. Therefore, the first step is to filter the data so that only the relevant data is left by identifying specific customer service “queues” that are part of the department under study. This data is then enriched by joining credit application volumes as well as financed agreement volumes and finally with temporal dimensions, most importantly the year and the week, to be able to have a variable for aggregating to weekly sums.

The data is then split into a training and testing set. The most recent 33 (20%) observations of the data act as the testing set and the first 133 (80%) as the training set. In date terms this means that data after the week starting 31 July 2023 is included in the testing set while everything up to and including the week starting 31 July 2023 comprises the training set.

3.3 Model selection

As explained in chapter 2.5, time-series analysis methods, more specifically ARIMA and exponential smoothing models, were chosen to be compared for model selection. This is due to several factors. First, the needed forecast is on a weekly level and, according to numerous studies, a limited dataset, as opposed to hundreds or thousands of datapoints in an intraday sample, does not often see benefits from more complex models such as neural nets or ensemble models (Bertsimas & Boussioux, 2023; Ibrahim et al., 2016; Oliveira et al., 2014; Saccani, 2013; Shmueli & Lichtendahl, 2016).

Second, the forecasting process inside the company needs to be simple enough to be maintained also in the future. A less complex forecast has been shown to be a good starting point since small prospective gains in model performance might not be worth risking the interpretability and ease of maintenance of a forecast model, since the most important thing is that the model is used and implemented in the first place (Cleveland, 2019; Saccani, 2013).

The ARIMA model is also tested both with and without external variables, agreement volume and application volume. All in all, this means that four different models are being compared to find the best fit.

3.3.1 Data exploration

Supported by the findings of the literature review and especially call center demand forecasting, the simpler to implement time series models of ARIMA and exponential smoothing (ETS, from Error Trend Seasonality) were chosen to be tested for this study, as no forecasting capabilities for the processes under focus currently exist at the case company, and more complexity seems not to necessarily bring extra benefits (Saccani, 2013) so introducing complexity for the sake of complexity should be avoided.

First, however an exploratory data analysis should be conducted, which includes the key summary statistics of the data under study, such as the number of data points, their variability and number of variables, as well as visualizations that make it easier to interpret the type of data being analyzed (Abbott, 2014). Understanding the data on hand, and its characteristics is an important part in selecting what models should be selected for analyzing the data (Andrews, 2021). Important characteristics to determine include ordering, distances, and continuity as well as aggregated statistics (Andrienko & Andrienko, 2006).

3.3.2 Selection criteria

The ARIMA and exponential smoothing models have numerous parameters that influence the accuracy of the forecast. Many software packages including Alteryx, the one used in this thesis as it is based on R programming, provide tools for identifying the parameters that lead to the best fit based on the training and testing data sets.

The ordering parameters of the autoregressive models were chosen with R programming *auto.arima* function, which compares all available combinations and selects the best one based on the Hyndman-Khandakar algorithm (Hyndman & Athanasopoulos, 2021). The algorithm selects the best fitting model from a variety of options based on unit root tests, corrected Akaike Information Criteria (AICc) values, and maximum likelihood estimates. The smoothing parameters of the ETS model were likewise chosen with R programming *ets* function, which selects the best parameters using AICc values (Hyndman & Athanasopoulos, 2021).

AICc is designed to provide an accurate measure of model quality in situations where the sample size is small. AICc adjusts the original AIC by incorporating a correction for finite sample sizes. This correction becomes particularly significant when the ratio of sample size to the number of estimated parameters is low. The formula for AICc includes an additional term that penalizes the complexity of the model more heavily as the sample size decreases. By balancing model fit and complexity more effectively than the original AIC, AICc is widely recommended for use in model selection to avoid overfitting (K. Burnham & Anderson, 2004).

Maximum likelihood estimation on the other hand works by defining a likelihood function based on observed data and the model parameters, transforming it into a log-likelihood function and then finding the parameter values that maximize this log-likelihood function (Casella & Berger, 2001). This method is widely used due to its desirable statistical properties and its applicability to a broad range of models.

When the best fitting parameters have been chosen, the four models are compared to each other based on multiple commonly used error terms that measure the models' forecasting accuracy, listed in table 3, that are part of the Alteryx model comparison tool. It is common to use several error terms to evaluate model performance for better results and the ones listed are among the most popular ones (Shmueli & Lichtendahl, 2016).

Table 2. Compared error terms.

Abbreviation	Meaning
ME	Mean error
RMSE	Root mean squared error
MAE	Mean absolute error
MPE	Mean percentage error
MAPE	Mean absolute percentage error
MASE	Mean absolute scaled error

According to Box and Jenkins, while trying to find the best possible forecast possible, it is also crucial to present the user with some notion of the model's accuracy as a form of risk control (Box et al., 2015). The main way to achieve this is by including probability limits, or confidence intervals, in the presentation of the forecast so that it is clear to the user how confident they can be that the forecasted values will be within a certain range.

4 Results

The results are split into two parts. First, an exploratory data analysis to determine the general characteristics of the data being analyzed. Second, the results of the time series forecasting analysis are presented and compared.

4.1 Exploratory data analysis

Before deciding on a model or models to be used as the basis of a data analysis task, it is important to know what the data looks like as covered in chapter 3.3.1.

As this thesis is dealing with temporal data, we can determine and see from a graphical representation in figure 10, that the data at hand is an ordered set (ordered by date) with a distance of seven days, and discrete, as we are dealing with specific time intervals, instead of a sliding scale as would be the case in for example temperature data. A normalized sample of the data can also be found in appendix 5.

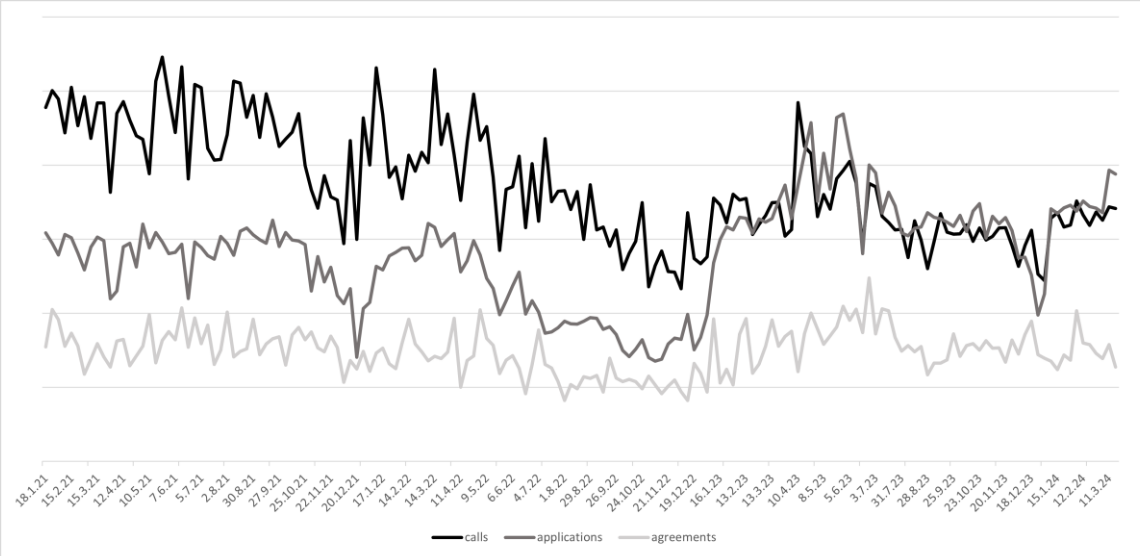

Figure 10. Source data time-series line graph including call, credit application, and financed agreement volumes collected from the case company's SQL database.

In table 2 we can see aggregated statistics such as average, median, maximum, and minimum values as well as standard deviation for each of the three variables under study for all 166 observations. The statistics have been normalized for easier comparison, so that the maximum value of each variable is equal to 1 and the others are presented as a fraction of that maximum. In other words, if the highest weekly number of calls is 2000, then a value of 1000 would be normalized to 0,500.

Table 3. Source data descriptive statistics, normalized values.

Variable	Average	Median	Max	Min	Std deviation	Observations
Calls	0,685	0,657	1,000	0,428	0,138	166
Applications	0,600	0,627	1,000	0,289	0,148	166
Agreements	0,601	0,600	1,000	0,333	0,123	166

From the aggregated statistics we can see that the variables are all relatively stable. Average and median do not have strong differences, signifying that the data population most likely does not have significant outliers in either direction. This is further supported by the minimum values, with a roughly equal distance to the average as the maximum values, as we can also see from the graphical representation in figure 11.

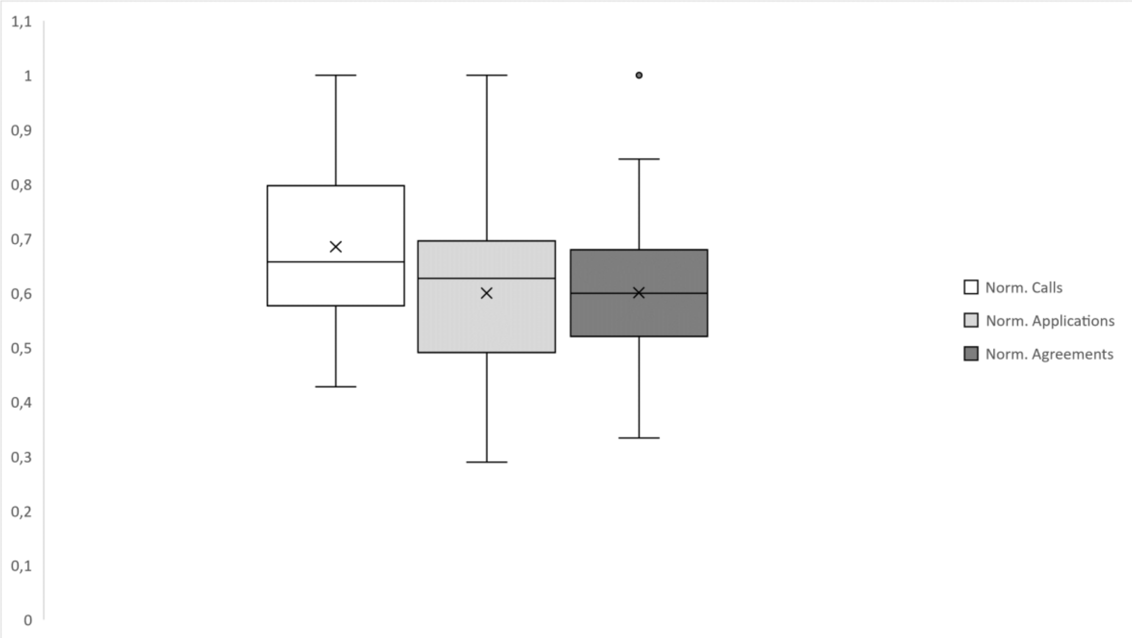

Figure 11. Box plot of weekly aggregated calls data, applications data, and agreements data.

4.2 Time series forecasting model comparison

Four time series forecasting models were chosen to be compared to find the best fit. These are an exponential smoothing model (ETS), an ARIMA model, and two ARIMAX models one using number of financed agreements as the explanatory variable, while the other using number of credit applications.

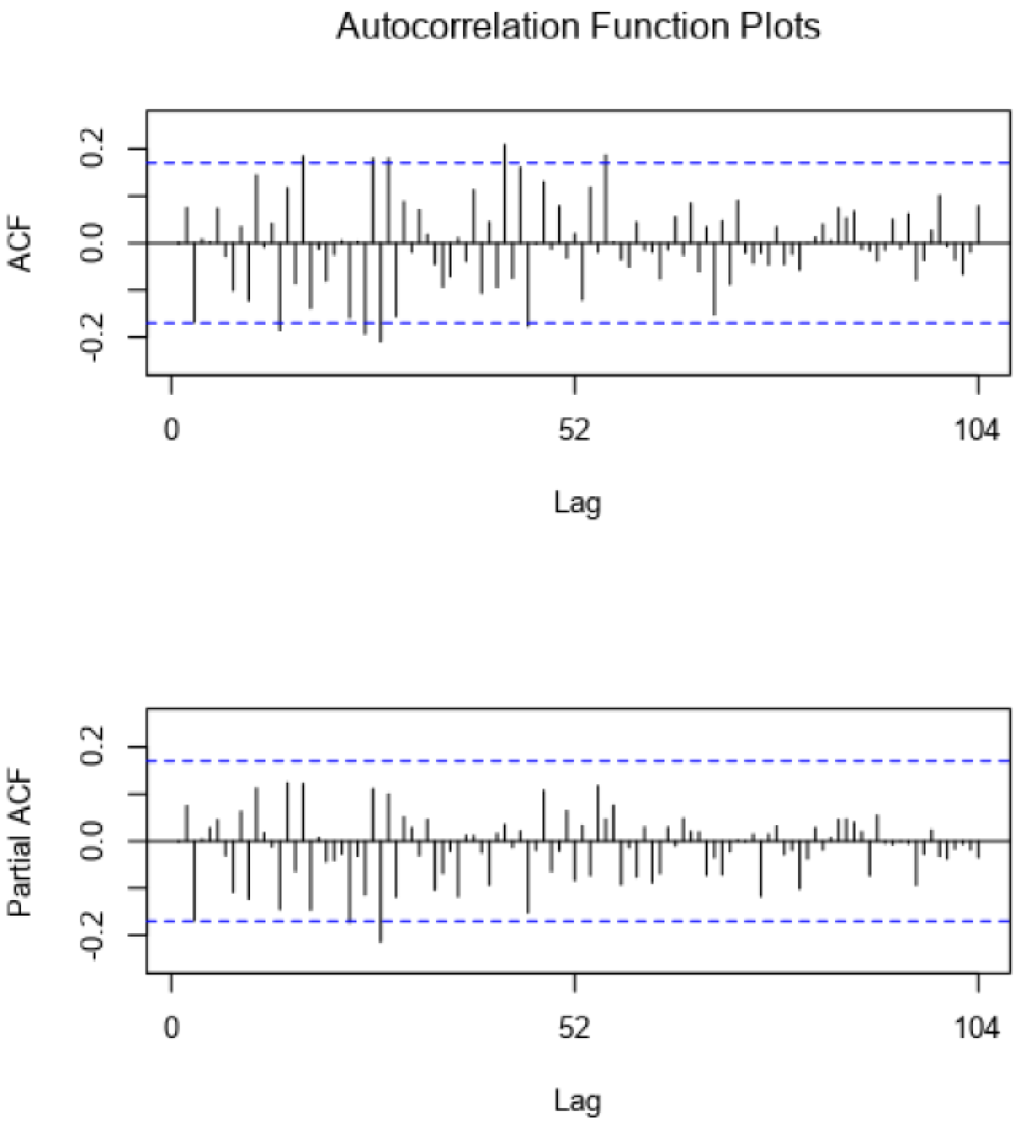

Figure 12. ACF and PACF plots of aggregated calls data.

The autocorrelation function, or ACF, as well as the partial ACF (PACF) for the aggregated calls data, the basis for the ARIMA model, are plotted in figure 12. The ACF and PACF can be used in determining the orders of the AR and MA components to be used, as they show the correlation between current and past values. The more specific model parameters and R functions for all compared models can be found in the appendix. The selection was made with a Hyndman-Khandakar algorithm explained in the methodology, based on the values of figure 12. As the plot is sinusoidal, we can detect some level of seasonality. The ACF plot is also developing towards zero, so we can determine that there is at least some level of autocorrelation.

Despite this, the best fitting ARIMA model without external variables was an ARIMA(0,1,1) model, which is very close to an exponential smoothing model. This would suggest that the amounts of autocorrelation are not significant enough to add value to the model, as a simpler one was chosen based on the unit root test, AICc score and MLE. The best fit suggests that the data shows a trend over time but observations are not heavily correlated with past values, while the model residuals of the previous observations influence the current period.

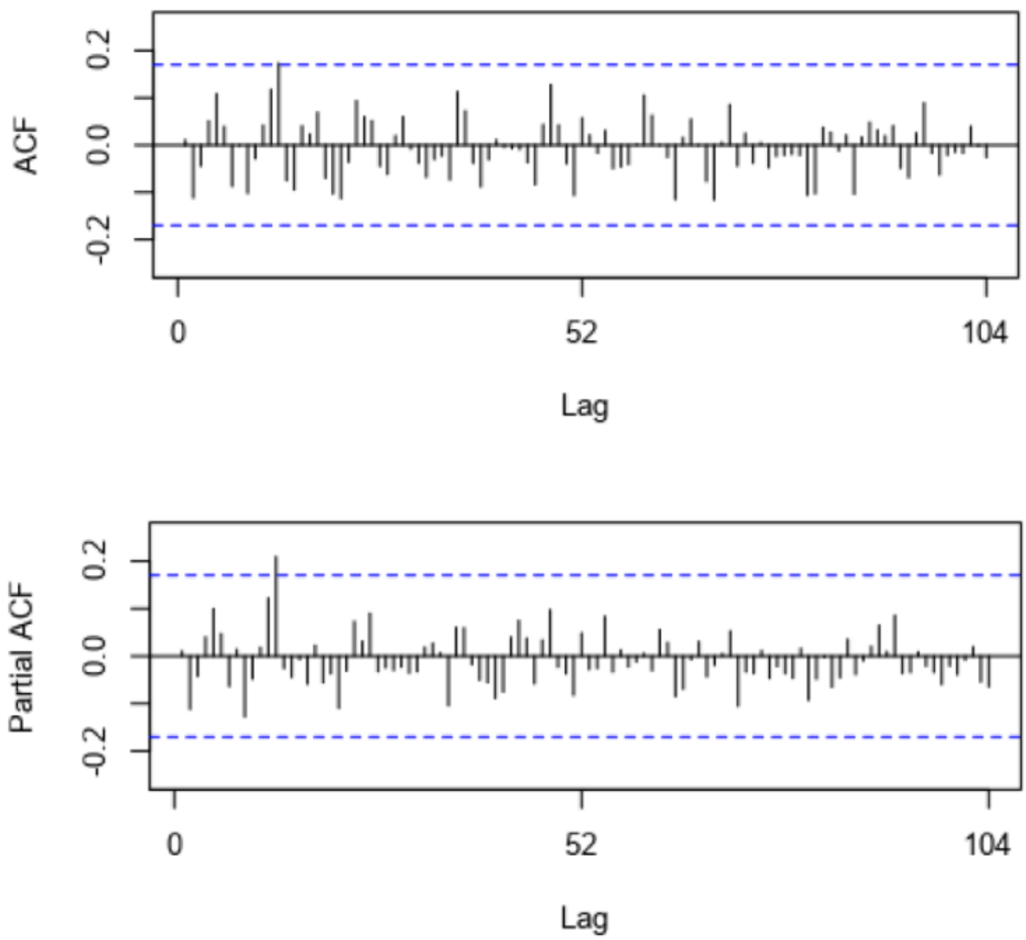

Figure 13. ACF and PACF plots of aggregated calls and application volumes.

The best fitting ARIMAX model including application volumes was ARIMA(1,0,0). This means that the model has one autoregressive term, but without differencing or moving average component, in other words a first-order autoregressive model. The observations are therefore best predicted by a relatively strong correlation to the previous value. The ACF and PACF plots for this model are described in figure 13.

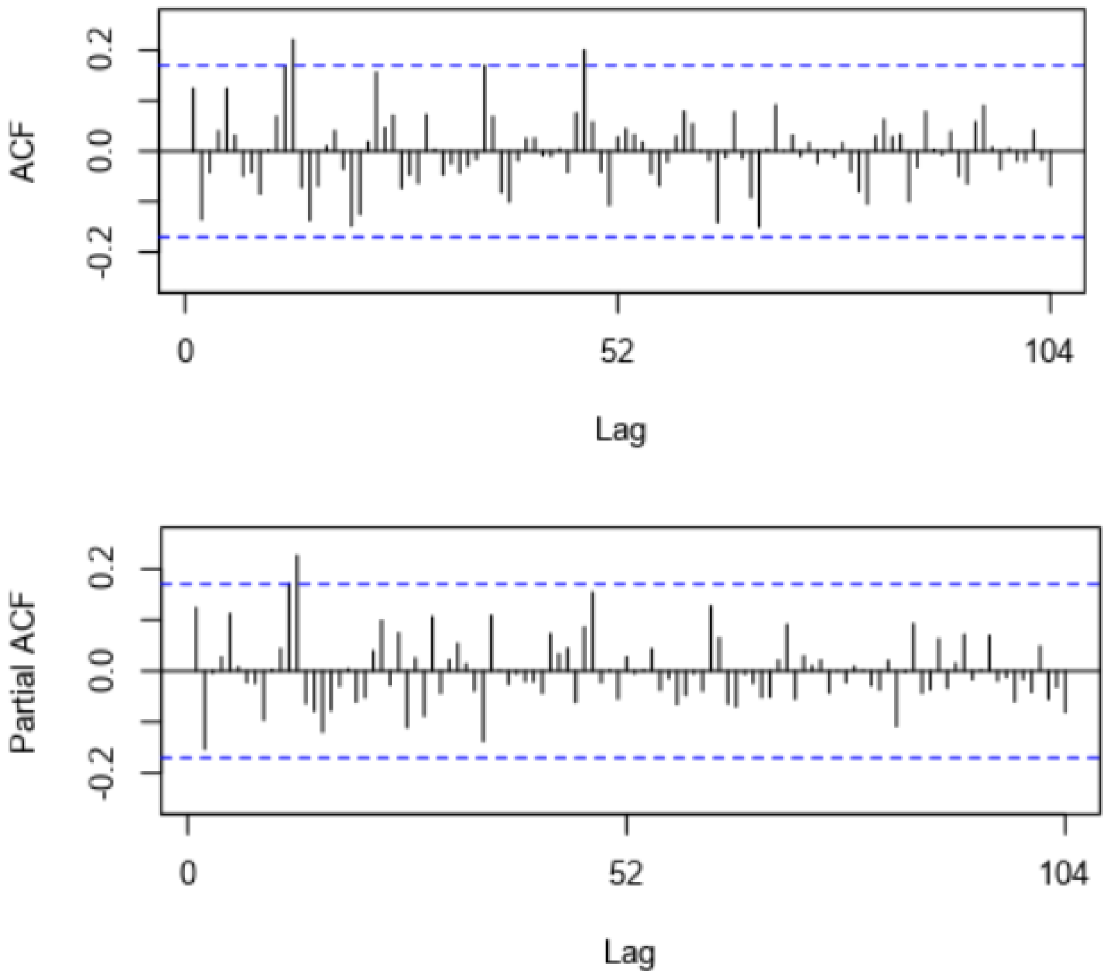

Figure 14. ACF and PACF plots of aggregated calls and agreement volumes.

For the ARIMAX model with agreement volumes, the best fitting parameters are again ARIMA(0,1,1) suggesting similar characteristics to the ARIMA model without external variables where the data exhibits some trend over time but observations are not very correlated and which includes one lagged error term. The ACF and PACF plots for this model are described in figure 14.

The chosen exponential smoothing model was an ETS(A,N,N) model. Figure 15 presents the decomposition components of the ETS model, in this case only the observed and level, or baseline, plots. No slope or seasonal plots were generated. The ANN model was chosen by comparing AICc values of different options and is also supported by comparison of our graphed data in figure 10 and Svetunkov's table in figure 6. The ANN model means that the model is a simple exponential smoothing model with additive errors where alpha is 0,278565. The alpha denotes the rate at which past observations are weighted as the smoothed values are calculated.

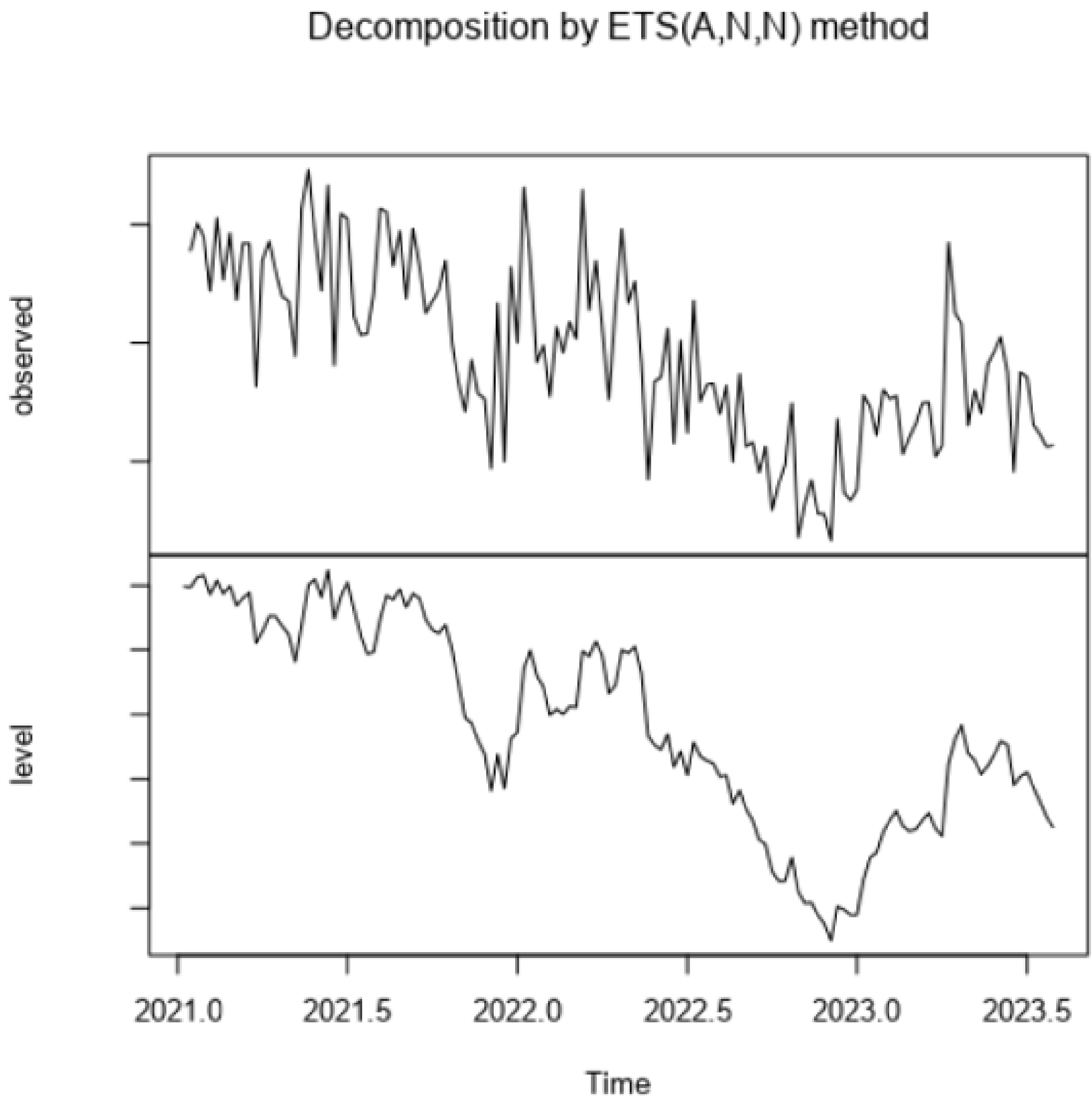

Figure 15. ETS(N,A,A) decomposition plot of aggregated calls data.

The models are then compared based on several error terms that in essence track how far the predicted values are from actual testing data on various metrics. These metrics are listed in table 3. We can see from table 4 that the ARIMA(0,1,1) model, here called "Arima_auto" has the lowest error term on every metric, while the ETS model is almost identical. Both are simple forecasts with flat predictions but slightly different parameters for estimating the moving average level, leading to a slight edge for the ARIMA model.

Table 4. Comparison of models' error terms.

Model	ME	RMSE	MAE	MPE	MAPE	MASE
ARIMA_auto	-105.334	167.1648	126.0849	-7.65	8.8592	0.5318
ETS_ANN	-107.8244	168.7451	127.5188	-7.8127	8.9596	0.5378
ARIMAX_creditapps	-511.2508	519.6229	511.2508	-33.8121	33.8121	1.196
ARIMAX_agrs	-413.3322	436.9219	413.3322	-27.6886	27.6886	0.9669

Figure 16 displays the plotted values of the compared forecasting models and we can see that the winning ARIMA model is almost equal to the chosen ETS configuration. The longer term downward trend seems not to be accounted for by the ARIMAX models tested, as there is no corresponding long term trend for the external variables, as can be seen from figure 10.

.

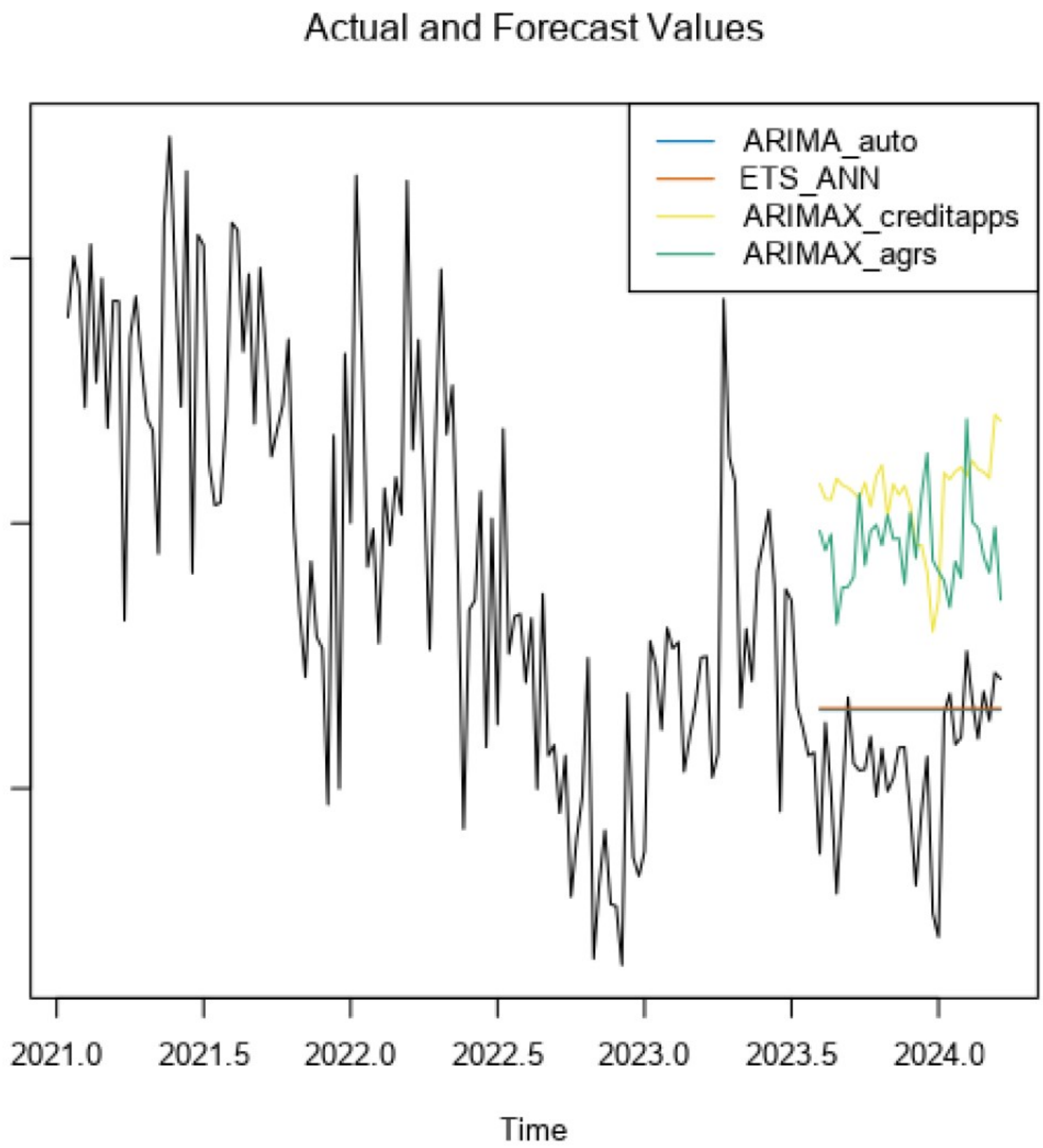

Figure 16. Plotted forecasts on top of training and testing data.

4.3 Study summary and contributions to the existing literature

The analysis begins with an examination of the temporal data, which is an ordered set with weekly intervals and discrete time points. Initial descriptive statistics indicate the data is relatively stable, with means and averages closely aligned, suggesting a lack of significant outliers. Four forecasting models were selected for comparison. An exponential smoothing model (ETS), an ARIMA model, and two ARIMAX models—one using the number of financed agreements and the other using the number of credit applications as explanatory variables.

Model comparisons based on several error metrics revealed that the ARIMA(0,1,1) model, referred to as "Arima_auto," had the lowest error terms across all metrics, slightly outperforming the ETS model. Both models provide simple, flat forecasts with slight differences in parameter estimation for the moving average level, giving the ARIMA model a slight edge in predictive accuracy.

The identified model can then be used to generate a forecast for the upcoming weeks, allowing the call center management to have some preliminary indication for whether call center volumes are trending up or down during the forecasted period. They are then able to use this indicative forecast to create a resourcing plan and make more informed staffing decisions. Another valuable finding based on the results is that it seems like new financing and application volumes in this financial services call center do not have a significant impact on call center volumes, at least in the 133 week time period that made up the training data set.

5 Conclusions and discussion

This final part of the thesis is divided into four sections. First, a discussion around the three identified research questions this thesis set out to answer is conducted. Then a description of how the finished model has been implemented into the case company's organization based on the findings in this thesis. Finally, the limitations of the research as well as recommendations for further subjects of study are presented.

5.1 Responses to the three identified research questions

This thesis set out to answer three questions, the first one concerning what types of models should be tested and compared when building an initial time series forecasting model for an organization which does not yet have a process in place. According to the literature review, even a simple model is an important tool in call center management (Bodin, 1999; Cleveland, 2019; Saccani, 2013) and favoring simplicity does not necessarily mean less accurate forecasts (Saccani, 2013). Due to these findings, simpler models commonly used in time series analysis, such as ARIMA and ETS, are favored (Antipov & Meade, 2002; Barrow, 2016; Bugarčić & Radojičić, 2020).

With the second research question the goal was to find out what the best model is out of the chosen ones when it comes to forecasting the call center data of the case company, in particular. An error term comparison was done between three types of ARIMA models, as well as an exponential smoothing model. The data was cleaned and aggregated to a weekly level containing 166 months of data and divided into a training and testing set. The parameters of the time series models were chosen using R programming tools to come up with the most accurate ones.

In the end, the simpler models were the most accurate, with an ARIMA model with no external variables having the smallest error terms across the board, followed by a simple exponential smoothing model with additive errors. This is the model that is used in building the forecast process.

The second question in essence also answers the third research question. The goal was to investigate whether the external variables identified in the organization, new agreement volumes and credit application volumes, were driving call center volumes to a significant extent. This seems not to be the case, at least to the extent that they would provide more accurate forecasts compared to just using the historical call center volumes on their own in building forecasts. The explanation for this might be that while new agreement and application generation obviously does generate customer contacts, the majority of the call volume is driven by the existing financed portfolio, instead of new agreements.

Another reason for the ARIMAX models not fitting the testing data as well as the simpler ARIMA model might be that there is a longer-term downward trend reflected in the call volume data, that is not present for either agreement or application volumes. This suggests that there might also be some process improvements, for example in the self-service channels available to customers, that have over time resulted in fewer contacts by customers, while the financing volumes have not been affected.

5.2 Implementation of the forecasting tool

With the most accurate model identified, it is possible to use the model to create a rolling forecast that is updated weekly, and which generates data for a specific number of weeks forward. In the case company, the need was for a forecast for the next four weeks. This data, alongside the historical development, is then visualized in a dashboard using the Tableau business intelligence software to be followed by the call center managers and operative employees in planning their near-term resource allocation.

Due to the features available in both software, the forecasting process as well as the updating of the dashboard are done completely automatically on the company's server. Due to not involving much actual coding, the process is also relatively simple to maintain in the future by the business intelligence team. The software also provides an 80% confidence interval for the forecast to enable variations to be visible to the users of the report. This is recommended as a form of risk control for end users in creating time series forecasts (Box et al., 2015).

5.3 Limitations of the study

From a data analysis perspective, the biggest limitations of the thesis are the sample size of 166 observations and, in a way, data availability. The limited number of observations, which is restricted by the case company's database data retention rules, and variables, which is limited by the data collection capabilities, essentially rule out some more sophisticated forecasting models, such as the use of neural networks.

As found in the literature review, one can achieve good results even without complex models, however there might be various factors that have a larger impact on call center volumes than their historical values apart from the ones included in this study. Testing these would require more extensive data collection in the organization. One clear limitation is that the subject of the calls is not recorded, so it is at the moment not possible to create accurate forecasts for specific customer issues, some of which take more time to handle than others.

Another limitation is that there are some external factors that have an effect on call center volumes that are hard to quantify in a statistical forecast. One such factor is the development of the company's self-service channels, which can quickly reduce the need for phone-based customer queries. This is something that must be taken into account by the call center managers when interpreting the forecast dashboard.

5.4 Further research

If data collection around call center activities at the case company becomes more extensive, it could be worthwhile to investigate what factors, if any, have a significant impact on volume changes. Data exists on the lengths of individual calls and if this could be combined with data points describing the subject of the call on any level of specificity, it would be possible to create even more granular forecasts down to the daily and hourly level. A daily or hourly forecast would also increase the number of data points available within the constraints of the company's data retention rules. This would potentially also allow for more flexibility methodologically. More sophisticated modeling techniques could also be used in building such a model, adding to the literature on ensemble models and neural networks in call center forecasting for example, as more data points would be available.

Finally, when it comes to operative forecasting there is another unexplored area within customer service with plenty of data points and a possibility for many types of analysis enabled by emerging pattern recognition capabilities. This area is emails, which are also tracked within a customer engagement software platform. With a view into how much time certain tasks take on average, there is much potential in having an even clearer picture of the company's future operative resource allocation.

6 References

Abbott, D. (2014). *Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst*. John Wiley & Sons, Inc.

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. https://doi.org/10.1111/j.1937-5956.2007.tb00288.x

Andrews, M. (2021). *Doing Data Science in R: An Introduction for Social Scientists* (1st Edition). SAGE Publications Ltd.

Andrews, R. (1989). *The Concise Columbia Dictionary of Quotations*. Columbia University Press.

Andrienko, N., & Andrienko, G. (2006). *Exploratory Analysis of Spatial and Temporal Data*. Springer-Verlag. https://doi.org/10.1007/3-540-31190-4

Anglou, F. Z., Ponis, S., & Spanos, A. (2021). A machine learning approach to enable bulk orders of critical spare-parts in the shipping industry. *Journal of Industrial Engineering and Management*, *14*(3), 604–621. https://doi.org/10.3926/jiem.3446

Antipov, A., & Meade, N. (2002). Forecasting call frequency at a financial services call centre. *Journal of the Operational Research Society*, *53*(9), 953–960. https://doi.org/10.1057/palgrave.jors.2601415

Barrow, D. K. (2016). Forecasting intraday call arrivals using the seasonal moving average method. *Journal of Business Research*, *69*(12), 6088–6096. https://doi.org/10.1016/j.jbusres.2016.06.016

Bertsimas, D., & Boussioux, L. (2023). *Ensemble Modeling for Time Series Forecasting: an Adaptive Robust Optimization Approach*. http://arxiv.org/abs/2304.04308

Bodin, M. (1999). *Maximizing Call Center Performance : 136 Innovative Ideas for Increasing Productivity and Customer Satisfaction* (1st Edition). CRC Press.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*, 5–59.

Box, G. E. P., Jenkins, G., Reinsel, G., & Ljung, G. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley.

Bugarčić, P., & Radojičić, V. (2020). Forecasting demand for a telephone call center. *INTERNATIONAL JOURNAL FOR TRAFFIC AND TRANSPORT ENGINEERING*, *10*(3), 266–277. https://doi.org/10.7708/ijtte.2020.10(3).01

Burnham, K., & Anderson, D. (2004). *Model Selection and Multimodel Inference* (K. P. Burnham & D. R. Anderson, Eds.). Springer New York. https://doi.org/10.1007/b97636

Casella, G., & Berger, R. (2001). *Statistical Inference* (2nd Edition). Cengage Learning.

Cleveland, B. (2019). *Contact Center Management on Fast Forward: Succeeding in the New Era of Customer Experience* (Fourth Edition). ICMI.

Cryer, J. D., & Chan, K.-S. (2010). *Time Series Analysis: With Applications in R* (Second Edition). Springer Texts in Statistics.

Ding, S., & Koole, G. (2022). Optimal call center forecasting and staffing. *Probability in the Engineering and Informational Sciences*, *36*(2), 254–263. https://doi.org/10.1017/S0269964820000595

Edna Chan, S., Ohlmann, J., Dunbar, S., Dunbar, C., Ryan, S., & Savory, P. (2002). Operations Research Methods Applied to Workflow in a Medical Records Department. *Health Care Management Science*, *5*(3), 191–199.

Fijorek, K., & Leśniewska, A. (2013). Statistical Forecasting of the Indicators of Polish Airport's Operations. *Folia Oeconomica Stetinensia*, *11*(1), 7. https://doi.org/10.2478/v10031-012-0010-0

Gans, N., Koole, G., & Mandelbaum, A. (2003). Telephone Call Centers: Tutorial, Review, and Research Prospects. In *Manufacturing and Service Operations Management* (Vol. 5, Issue 2, pp. 79–141). INFORMS Inst.for Operations

Res.and the Management Sciences. https://doi.org/10.1287/msom.5.2.79.16071

Gans, N., Shen, H., Zhou, Y. P., Korolev, N., McCord, A., & Ristock, H. (2015). Parametric forecasting and stochastic programming models for call-center workforce scheduling. *Manufacturing and Service Operations Management*, *17*(4), 571–588. https://doi.org/10.1287/msom.2015.0546

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Hatzakis, E. D., Nair, S. K., & Pinedo, M. L. (2010). Operations in Financial Services-An Overview. *Production and Operations Management*, *19*(6), 633–664.

Hawkins, D. M. (2004). The Problem of Overfitting. In *Journal of Chemical Information and Computer Sciences* (Vol. 44, Issue 1, pp. 1–12). https://doi.org/10.1021/ci0342472

Hyndman, R., Ahmed, R. A., Athanasopoulos, G., & Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational Statistics and Data Analysis*, *55*(9), 2579–2589. https://doi.org/10.1016/j.csda.2011.03.006

Hyndman, R., & Athanasopoulos, G. (2021). *Forecasting: principles and practice* (3rd edition). OTexts.

Hyndman, R., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. In *International Journal of Forecasting* (Vol. 18). www.elsevier.com/locate/ijforecast

Hyndman, R., Koehler, A., Ord, K., & Snyder, R. (2008). *Forecasting with Exponential Smoothing*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-71918-2

Ibrahim, R., Ye, H., L'Ecuyer, P., & Shen, H. (2016). Modeling and forecasting call center arrivals: A literature survey and a case study. *International Journal of Forecasting*, *32*(3), 865–874. https://doi.org/10.1016/j.ijforecast.2015.11.012

Jakobs, M., & Saadallah, A. (2023). Explainable Adaptive Tree-based Model Selection for Time-Series Forecasting. *2023 IEEE International Conference on*

Data Mining (ICDM), 180–189. https://doi.org/10.1109/ICDM58522.2023.00027

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning* (Vol. 103). Springer New York. https://doi.org/10.1007/978-1-4614-7138-7

Kuhn, M., & Johnson, K. (2016). *Applied Predictive Modeling*. Springer.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. In *Journal of Forecasting* (Vol. 1).

Makridakis, S., Athanasopoulos, G., & Hyndman, R. (1997). *Forecasting: Methods and Applications* (3rd ed.). Wiley.

Mehrotra, V., & Fama, J. (2003). Call center simulation modeling: Methods, challenges, and opportunities. *Winter Simulation Conference Proceedings*, *1*, 135–143. https://doi.org/10.1109/wsc.2003.1261416

Oliveira, M., Phung, D., & Li, H. (2014). *Ensembles for Time Series Forecasting* (D. Phung & H. Li, Eds.; Vol. 39, pp. 360–370). JMLR: Workshop and Conference Proceedings.

Piccialli, F., Giampaolo, F., Salvi, A., & Cuomo, S. (2021). A robust ensemble technique in forecasting workload of local healthcare departments. *Neurocomputing*, *444*, 69–78. https://doi.org/10.1016/j.neucom.2020.02.138

Rausch, T. M., Albrecht, T., & Baier, D. (2022). Beyond the beaten paths of forecasting call center arrivals: on the use of dynamic harmonic regression with predictor variables. *Journal of Business Economics*, *92*(4), 675–706. https://doi.org/10.1007/s11573-021-01075-4

Rexer, K. (2020). *2020 Data Science Survey*. https://www.rexeranalytics.com/data-science-survey

Saccani, N. (2013). Forecasting for capacity management in call centres: Combining methods, organization, people and technology. *IMA Journal of*

Management Mathematics, *24*(2), 189–207. https://doi.org/10.1093/imaman/dpr027

Shmueli, G., & Lichtendahl, K. C. (2016). *Practical time series forecasting with R : a hands-on guide* (Second edition). Axelrod Schnall Publishers.

Svetunkov, I. (2023). *Forecasting and Analytics with the Augmented Dynamic Adaptive Model (ADAM)*. Chapman and Hall/CRC. https://doi.org/10.1201/9781003452652

Teoh Yi Zhe, I., & Keikhosrokiani, P. (2021). Knowledge workers mental workload prediction using optimised ELANFIS. *Applied Intelligence*, *51*(4), 2406–2430. https://doi.org/10.1007/s10489-020-01928-5

van Gent, P., Melman, T., Farah, H., van Nes, N., & van Arem, B. (2018). Multi-Level Driver Workload Prediction using Machine Learning and Off-the-Shelf Sensors. *Transportation Research Record*, *2672*(37), 141–152. https://doi.org/10.1177/0361198118790372

van Gils, T., Ramaekers, K., Caris, A., & Cools, M. (2017). The use of time series forecasting in zone order picking systems to predict order pickers' workload. *International Journal of Production Research*, *55*(21), 6380–6393. https://doi.org/10.1080/00207543.2016.1216659

Appendices

Appendix 1. ARIMA model parameters

Method: ARIMA(0,1,1)(0,0,0)

Call:
auto.arima(Calls_total)

Coefficients:

	ma1
Value	-0.715127
Std Err	0.065418

sigma^2 estimated as 63058.2471: log likelihood = -916.57578

Information Criteria:

AIC	AICc	BIC
1837.1516	1837.2446	1842.9172

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-20.1030974	249.218785	192.5347833	-2.4445265	10.1761739	0.4240508	-0.0019719

Ljung-Box test of the model residuals:
Chi-squared = 34.5437, df = 23, p-value = 0.057694

Appendix 2. ARIMAX model parameters (applications)

Method: ARIMA(1,0,0)(0,0,0)

Call:
auto.arima(Calls_total, xreg = covars, max.p = 2, max.q = 2, max.P = 1, max.Q = 1, ic = "aicc", allowdrift = TRUE)

Coefficients:

	ar1	intercept	Credits
Value	0.101258	1383.293665	0.417681
Std Err	0.087535	114.465297	0.080702

sigma^2 estimated as 119878.69005: log likelihood = -964.87347

Information Criteria:

AIC	AICc	BIC
1937.7469	1938.0594	1949.3083

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.4235229	342.3078431	292.9251148	-3.1694882	15.5301399	0.7163497	0.0105825

Ljung-Box test of the model residuals:
Chi-squared = 23.7621, df = 23, p-value = 0.417076

Appendix 3. ARIMAX model parameters (agreements)

Method: ARIMA(0,1,1)(0,0,0)

Call:
Arima(Calls_total, xreg = covars, order = c(0, 1, 1), seasonal = list(order = c(0, 0, 0), period = 52))

Coefficients:

	ma1	Agrs_total
Value	-0.999996	0.893899
Std Err	0.036159	0.183311

sigma^2 estimated as 121913.47073: log likelihood = -961.66755

Information Criteria:

AIC	AICc	BIC
1929.3351	1929.5226	1937.9835

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
37.0974287	345.2007283	287.9889664	-1.225702	15.1793924	0.7042783	0.1240043

Ljung-Box test of the model residuals:
Chi-squared = 36.3874, df = 23, p-value = 0.037646

Appendix 4. ETS model parameters

Method:
ETS(A,N,N)

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-20.1128424	249.209964	192.9101836	-2.453918	10.1980568	0.4248776	0.0038252

Information criteria:

AIC	AICc	BIC
2124.2831	2124.4692	2132.9542

Smoothing parameters:

Parameter	Value
alpha	0.278565

Initial states:

State	Value
l	2396.169084

Appendix 5. Normalized data sample

First 20 rows of aggregated weekly agreements, calls, and credit application data, normalized on a zero to one scale.

Record	OriginalDateTime	norm_agrs	norm_calls	norm_credits
1	18.1.2021 0:00	0,623890234	0,875412239	0,658422175
2	25.1.2021 0:00	0,827280065	0,917552217	0,628571429
3	1.2.2021 0:00	0,772397094	0,896665445	0,594882729
4	8.2.2021 0:00	0,628732849	0,813118358	0,653304904
5	15.2.2021 0:00	0,698950767	0,925613778	0,643923241
6	22.2.2021 0:00	0,629539952	0,830340784	0,600426439
7	1.3.2021 0:00	0,475383374	0,902161964	0,551812367
8	8.3.2021 0:00	0,557707829	0,798827409	0,617057569
9	15.3.2021 0:00	0,642453592	0,886771711	0,646055437
10	22.3.2021 0:00	0,569814366	0,886405277	0,635820896
11	29.3.2021 0:00	0,514931396	0,665811653	0,468230277
12	5.4.2021 0:00	0,655367232	0,860754855	0,490405117
13	12.4.2021 0:00	0,664245359	0,889703188	0,617910448
14	19.4.2021 0:00	0,522195319	0,843898864	0,62771855
15	26.4.2021 0:00	0,575464084	0,805423232	0,559914712
16	3.5.2021 0:00	0,628732849	0,796995236	0,68315565
17	10.5.2021 0:00	0,799031477	0,711982411	0,615351812
18	17.5.2021 0:00	0,538337369	0,940637596	0,659275053
19	24.5.2021 0:00	0,660209847	1	0,631982942
20	31.5.2021 0:00	0,707828894	0,902894833	0,597867804