Can we forecast lung cancer data accurately?

A Smith

October 2020

Executive summary

Using patient referral/test data for lung cancer from hospitals in the Cwm Taf Morgannwg University Health Board (CTMUHB) in South Wales, this project examines the extent to which we can accurately forecast future hospital referrals/tests using statistical analysis methods. Wales lags other countries for lung cancer outcomes, having the second-worst survival rates for the disease in Europe. The CTMUHB region has the highest occurrences of the disease per capita within Wales. As a result of this poor survival rate, NHS Wales has introduced the National Optimal Lung Cancer Pathway to expediate a patient's journey through the healthcare system, ensuring early diagnosis/treatment and improved outcomes. To ensure effective planning and improve efficiency, it is important to offer healthcare providers some estimate of the demand they may expect on services.

Our project stems from data of the details of patient interaction with the lung cancer pathway which contain the number of referrals/tests requested per day/week/month across the CTMUHB. We transform the data into a series of datapoints ordered by time, a *time series*, and perform rudimentary analysis. We also segment the data by hospital and test type to examine these categories independently. This initial analysis concludes that the Prince Charles and Royal Glamorgan hospitals receive the bulk of referrals, and that CT and PET tests are the most common tests undertaken. We create boxplots to determine if we can detect patterns in the frequency of tests/referrals by day/month. We found that the numbers of tests/referrals are largest during the working week, and that referrals and tests are more prevalent in spring/summer than winter. We continue our analysis by discussing *moving averages*, which smooth the data and provide insight into underlying patterns. Next, we perform statistical tests to determine whether the data is stationary (does not depend on time), or nonstationary (displays *seasonality*, a regular and predictable change, or a long-term *trend*). We find statistical evidence of nonstationarity on the monthly *referrals* and weekly *tests*. Finally, we decompose each series to examine the seasonality, trend, and *random error*.

We partition each series into a training set (to train each model) and a test set (to test its accuracy). We apply statistical forecasting methods to predict future values past the end of the existing series. The statistical methods used increased in complexity and are listed in order of complexity below:

1. ***Naïve methods***: which predict that last observed value will continue unchanged.

2. ***Extrapolation models:*** which apply a mathematical function to all preceding data points to forecast future datapoints and trends.
3. ***Causal models****:* which assume a linear relationship (line of best fit).
4. ***ARIMA models****:* which include parameters for seasonality and trend.
5. ***Singular spectrum analysis:*** which deconstructs the series into components to reduce *noise*.
6. ***Artificial neural network (ANN):*** a method inspired by the workings of the brain.

The accuracy of the forecasting models is assessed using several different *error statistics*, each provides different information about how well our forecasting models fit the data. We fit each model to each series and calculate the forecasts made and the error statistics produced. After analysis of each method, we endorse one method for each time series, based on its error statistics and a plot of its efficacy. We determine that ANNs are the method most often endorsed, though they may be unsuitable for linearly structured data. We find that a type of extrapolation model, *single exponential smoothing* (SES), is favourable against ANNs for non-noisy datasets. ANNs were effective on the noisy *daily* data, however other methods may be preferred for less variable *monthly* data.

Table 0.1: frequency table- most preferred forecast method.

Forecasting method	Frequency
Artificial neural network	6
Single exponential smoothing	3
Holt-Winters (extrapolation model)	2
Simple linear regression (causal)	2
Singular spectrum analysis	2
ARIMA	1
Holt-Linear (extrapolation model)	1
Multiple linear regression (causal)	1
Total	**18**

Overall, ANNs scored the lowest error statistics, suggesting that they are usually the most accurate method. However, while ANNs are an exciting and effective method of forecasting, they are imperfect, and a statistician should consider the nature of the time series, as a different model may outperform an ANN. In some cases ANNs were prone to overfit the data and for noncomplex linear data, SES or causal models were more likely to be endorsed.

Finally, while time series forecasting methods are often extremely accurate, they may fail to account for extreme events and crises. We consider that our forecasts in the period of interest may be drastically altered by the impact of the COVID-19 pandemic, which radically altered medical practice and hospital capacity in early 2020.

Acknowledgements

My deepest gratitude to my parents for encouraging me to reach my potential. Your support during tough times has been hugely appreciated. I could have not progressed this far without your guidance and wisdom. Also to Charlie, who was always happy to accompany me late into the evening.

Finally, my thanks to my partner Kelly for all her love and support through distance.

Contents Page

Page

List of figures

Page

List of tables

Please note that values in tables are rounded to 2 decimal places, excluding whole numbers.

Page

1. Report

1.1 Introduction

Lung cancer is among the most diagnosed cancers worldwide, and a leading cause of morbidity across Europe[1]. In the UK, between 2015-2017, there were around 47,800 new cases of lung cancer per year[2]. Furthermore, most cases of lung cancer in the UK were diagnosed at the 'late' stage when the disease is least treatable. Early detection, diagnosis and treatment are crucial in the fight against lung cancer, therefore mathematical and statistical models which forecast lung cancer incidences should be considered to ensure that healthcare bodies can pre-empt demand for respiratory services.

This research examines lung cancer patients in Cwm Taf Morgannwg University Health Board (CTMUHB) in South Wales. By fitting various models to time series detailing the frequency of lung cancer referrals/tests, we assess the accuracy with which methods of time series forecasting can accurately predict demand on the healthcare system. Our research consists of examinations of the structure of each series, a discussion of forecasting methods and finally an experiment to decide which model should be used to forecast future demand. We begin with crude methods, such as *naïve models*, and examine how accurately these models perform. We progress to *extrapolation models*, which analyse time series to examine phenomena such as trend and seasonality. Finally, we discuss more complex methods: *ARIMA*s, *singular spectrum analysis* (SSA) and *artificial neural networks (*ANNs*)*.

Our dataset consists of patients who visited the CTMUHB, the dates of their visit and which services they used. The data spans from 2017-2019, therefore our predictive models will forecast demand from 2019-2020. Where appropriate, data is grouped into daily/weekly/monthly figures. We focus on forecasting short-term data, from a timespan of around 10 days to 3 months. Our research will be valuable in determining the extent to which demand on the healthcare system can be effectively modelled, and how accurately we can forecast such data. Should our models prove effective, they may be adopted by NHS Wales to ensure that facilities are prepared for demand on their services. Furthermore, if the frequency of referrals/tests is much lower than the figures forecasted by our models, it may prompt investigation into why referrals/tests are not occurring at the expected rate.

We test all series against each method, examining the impact of a series' structure on the endorsed forecasting method. After partitioning each series into a training set (to train each

model) and a test set (to assess each model), we discuss error statistics which determine the efficacy of each model, based on residuals. Calculating error statistics for each method allows us to determine the best performing model. Finally, we perform the preferred forecasting method on the entire series, predicting beyond the timespan of the dataset. We summarise our research objectives:

- Analyse the structure of each referrals/tests time series object, determining statistical information about each dataset.
- Discuss many forecasting techniques, from simple models to computationally complex machine learning models.
- Perform each technique on each series, calculating error statistics to determine which models provide the best fit.
- Forecast past the dataset timespan to produce future predictions.

1.2 Literature review

For the literature review, we examine 8 academic papers on lung cancer, found from both the scopus.com search engine and the NHS Wales website. Search criteria included papers on lung cancer in general, time series forecasting and research on forecasting cancer rates.

Lung cancer

Lung cancer is the 3rd most common type of cancer in the UK; however, it is the deadliest. The disease kills 35,000 people in the UK each year due to delays in diagnosis, treatment, and care (UK Lung Cancer Coalition 2019)[3]. Survival rates for lung cancer lag rates for all cancers: 15% compared with the 54% associated with all cancers. Lung cancer survival rates in the UK are much lower than other economically developed countries (UK Lung Cancer Coalition 2019)[3].

In Wales, lung cancer is also the deadliest cancer type. A UK Lung Cancer Coalition study found that the five-year survival rate for the disease within Wales was 6.6% between 2002-2006 (Lung Cancer Coalition 2016)[4], ranking Wales 28th out of 29 European countries for lung cancer outcomes. Only 12% of Welsh patients were diagnosed within the early stage of the disease, when it is most treatable. Initiatives have been implemented to improve lung cancer outcomes within Wales, including the 'Be Clear on Cancer' campaign from NHS Wales in 2016[4]. That year, a study from the Welsh Government and Macmillan Cancer

Support attempted to gain an insight into people's experience with cancer. Studies show extreme regional variation within Wales: between 33.3%-87.5% of patients receiving chemotherapy in 2014[4]. We consider a predictive model factoring this regional variation.

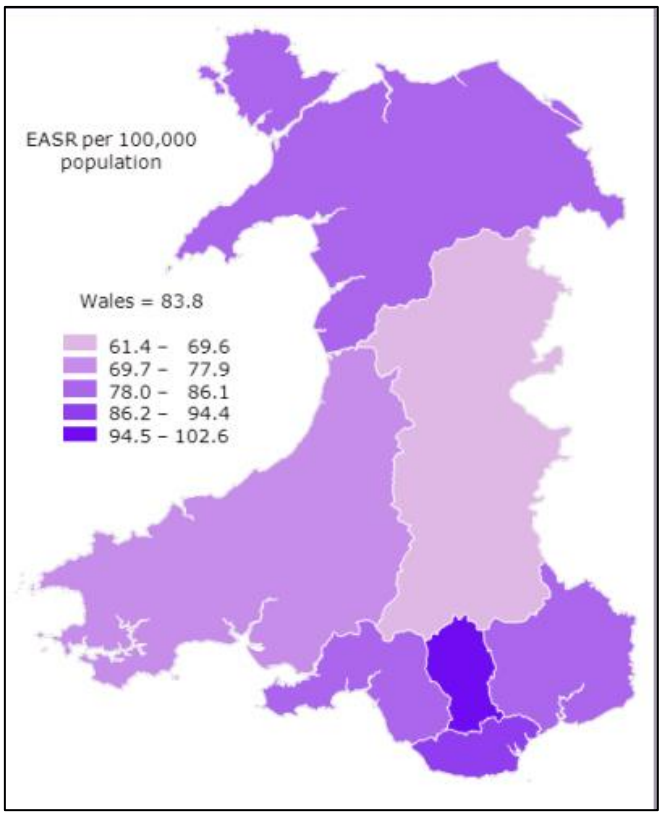

Figure 1.1: variation in lung cancer incidence rates across Wales[5].

Rates are highest in the CTMUHB region in South Wales and lowest in the Powys Teaching Health Board (PTHB) in Mid Wales[5].

National Optimal Lung Cancer Pathway (NOLCP)

Given the poor outcome for UK lung cancer patients, the NOLCP was introduced in August 2017 to expediate the patient journey. Delays in lung cancer treatment are associated with poor outcomes; the NOLCP was introduced to decrease time from referral to treatment (from 62 days to 49 days) (UK Lung Cancer Coalition 2019)[6]. A UK Lung Cancer Coalition report recommends using economic modelling to assess the impact of the pathway on turnaround times, patient experience, and service improvements (UK Lung Cancer Coalition 2019)[6].

This modelling should include predictions of demand on local hospitals, ensuring they have forecasts for demand on referrals/tests. The pathway can be seen below:

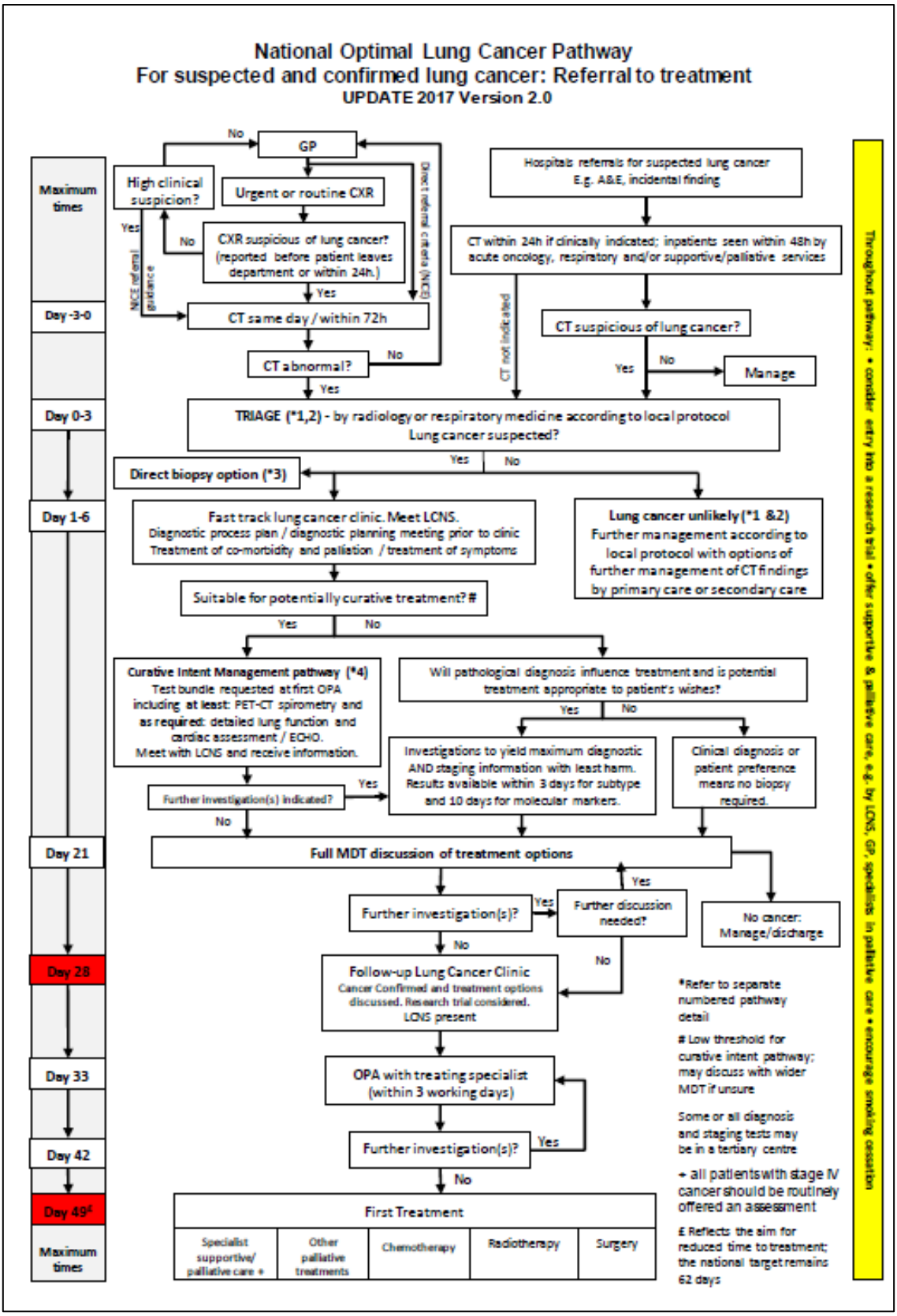

Figure 1.2: the NOLCP[7].

NHS Wales implemented their pathway, the National Optimal Pathway (NOP) for lung cancer patients in August 2017, aiming to improve outcomes, reduce regional variation and standardise care across the country (Welsh Thoracic Oncology Group 2019)[8]. The NOP expediated patients' journey through the healthcare system (Welsh Thoracic Oncology Group 2019)[8]. The pathway is a timed decision tree, mandating that a patient's MDT (multidisciplinary team) discussion occurs within 28 days and first treatment occurs within 49 days. A predictive model would provide an estimate of time between referrals/tests.

Forecasting and time series

We define a *time series* as a series of datapoints organised in time order and consider forecasting methods to predict *referrals*/*tests* time series across Welsh hospitals. Traditional forecasting methods consist of extrapolation models such as *single exponential smoothing* (SES) and *Holt* models (Madden & Tan 2006)[9]. For telecommunications data, these models outperform baseline *naïve* models which assume that current values continue indefinitely. SES and Holt models also outperformed *random walk* models which randomise the naïve model. SES and Holt models are also robust on imperfect and short datasets (Madden & Tan 2006)[9]. As our datasets are short, these techniques are promising. Expanding on previous research, we assess the accuracy of telecommunications forecasting methods on healthcare data.

More complex forecasting models, including *autoregressive integrated moving averages* (ARIMAs) and ANNs have also been devised. ARIMA models incorporate many exponential smoothing models (such as SES) and are demonstrably effective (Zhang 2002)[10]. Recently, *machine learning* (ML) techniques such as ANNs have been adapted for time series forecasting. ANNs modify themselves based on the structure of the data and are suitable for nonlinear modelling. Zhang suggests no model definitively always outperforms all others; hybrid models should be considered for complex problems with linear/nonlinear structures.

Forecasting cancer data

ARIMA models have generated forecasts of cancer incidence in Kenya (Langat, Orwa & Koima 2017)[11]. These models are trained on 70-80% of known data and assessed on the remaining 20-30% data. Langat, Orwa & Koima analysed cancer incidence rates across Kenya from 2000-2018. They found that ARIMA modelling is effective when forecasting incidences of cancer, producing a low *mean absolute percentage error* (MAPE) of 6.26. The

Kenyan study examined yearly data; we expand on this research by assessing the effectiveness of ARIMA models on daily/weekly/monthly data.

ANN models often score a lower MAPE than ARIMAs (Ezhil & Vijayalakshmi 2012)[12]. For 50 years of American bowel cancer incidence data, ANN models scored a MAPE of around 1.3 whereas ARIMA models scored around 2.7. Here, ANN models are favoured in modelling cancer data, particularly as the data displayed a non-linear structure. The research advocates the use of ANNs in forecasting bowel cancer incidence. We expand this research, performing similar analysis on our lung cancer data to examine whether ANN models outperform ARIMA models.

1.3 Analysis

Here we discuss the dataset which forms the basis of our analysis. After preparing the data, we perform some rudimentary statistical analysis to extract key information about its structure. Such analysis includes summary statistics, boxplots, moving averages, stationary and decomposition.

1.3.1 Data preparation

We examine data comprising of an excel spreadsheet detailing lung cancer patients' interactions with the CTMUHB; figure 1.3 displays a subset. The spreadsheet is split into 3 tabs: *maindata, testdata* and *explanation*. *Maindata* contains 1432 rows (one for each patient). The columns of the spreadsheet represent details relating to the patient's treatment: care facility, referral source, date of suspicion, date referral received, date of patient's first

appointment, date of diagnosis etc. We examine hospital referrals, referral sources and suspicion dates.

	A	D	K	L	M	Q
1	LinkID	CareFacility	DateSuspicion	DateReceived	DateFirstAppt	DateClosed
2	4689	Royal Glamorgan Hospital	02/01/2018	02/01/2018		16/01/2018
3	126M801205321	Royal Glamorgan Hospital	02/01/2018	03/01/2018		09/01/2018
4	5299	Royal Glamorgan Hospital	03/01/2018	11/02/2018		23/05/2018
5	4593	Royal Glamorgan Hospital	03/01/2018	03/01/2018		15/02/2018
6	126M801214401	Prince Charles Hospital	03/01/2018	05/01/2018		08/01/2018
7	126M801204042	Royal Glamorgan Hospital	03/01/2018	03/01/2018	10/01/2018	03/01/2018
8	4597	Royal Glamorgan Hospital	04/01/2018	04/01/2018		24/04/2018
9	5401	Royal Glamorgan Hospital	05/01/2018	05/01/2018		25/05/2018
10	126M801218283	Prince Charles Hospital	06/01/2018	08/01/2018	10/01/2018	10/01/2018
11	4685	Royal Glamorgan Hospital	07/01/2018	07/01/2018		26/01/2018
12	126M801221814	Royal Glamorgan Hospital	09/01/2018	09/01/2018		17/01/2018

Figure 1.3: subset- MainData tab.

	A	B	C	D	E
1	LinkID	TestName	TestRequestedDate	TestPerformedDate	TestReportedDate
2	1000001258	EBUS	25/10/2019	26/11/2019	26/12/2019
3	126M606183576	MRI	24/12/2019	24/12/2019	24/12/2019
4	1000001258	PET	11/12/2019	19/12/2019	19/12/2019
5	126M601838081	LFT	12/12/2018	19/12/2019	19/12/2019
6	126M606239724	CT Guided Biopsy	11/12/2019	17/12/2019	17/12/2019
7	126M803003322	CT Guided Biopsy	11/12/2019	16/12/2019	16/12/2019
8	126M606234649	X-ray	13/12/2019	13/12/2019	13/12/2019
9	126M606145056	CT Guided Biopsy	04/12/2019	12/12/2019	12/12/2019
10	126M606145056	CT Guided Biopsy	04/12/2019	12/12/2019	12/12/2019
11	126M606233539	CT chest	09/12/2019	12/12/2019	12/12/2019
12	126M601838081	CT chest	12/12/2018	12/12/2019	12/12/2019
13	126M606183576	PET	04/12/2019	04/12/2019	04/12/2019
14	126M606194705	CT TAP	15/11/2019	26/11/2019	04/12/2019
15	10936	MRI	02/12/2019	03/12/2019	03/12/2019
16	10936	CT TAP	02/12/2019	02/12/2019	02/12/2019
17	126M606196053	X-ray	02/12/2019	02/12/2019	02/12/2019
18	126M606196053	CT Guided Biopsy	28/11/2019	02/12/2019	02/12/2019

Figure 1.4: subset- TestData tab.

Figure 1.4 contains details of diagnostic tests performed by hospitals for each patient (*testdata* tab). This *tests* dataset is larger at 1944 rows as many patients received multiple tests. The columns represent details of each test: date requested, performed, and reported. We

merge patient ID numbers from tab 1 into the tests tab, observing which hospital performed each test.

There were issues to address in the *tests* dataset:

1. Outliers: the earliest observation for *DateRequested* occurs 11 months before the rest of the data, skewing plots and altering our data's structure. This may be a data entry error; we remove this datapoint.
2. Incomplete data: many datapoints are missing a *DatePerformed* or *DateReported*.

We aggregate the data and extract frequencies of referrals/tests for different time periods, using Excel to determine which day/week/month each test was performed. Next, we tabulate our main dataset, determining tests and referrals numbers for each day/week/month.

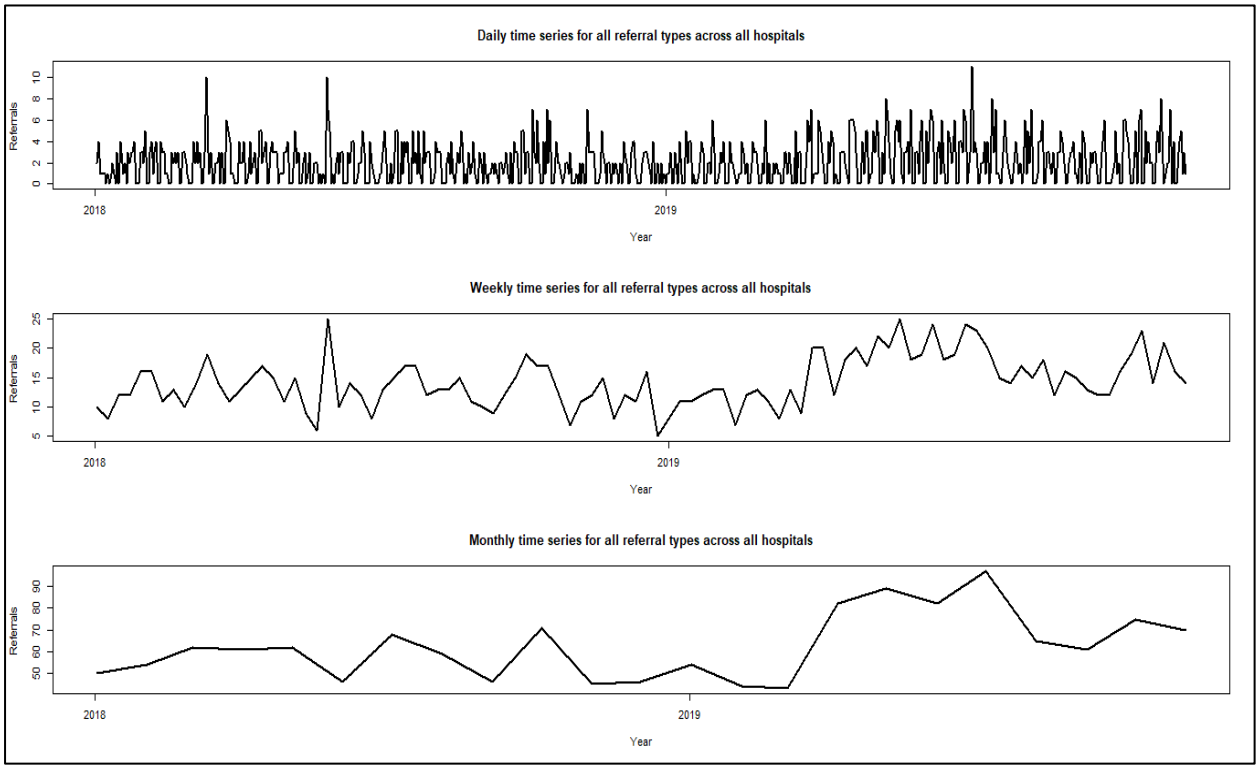

Figure 1.5: daily/weekly/monthly referrals (all hospitals).

The referral numbers fluctuate by day. Often there are zero referrals, however there may be as many as 10. The bottom two graphs (weeks/months) are less noisy. There is one busy week around mid-2018; the referral numbers increase towards the end of the period. We partition the data by hospital and referral/test type:

Table 1.1: frequency table- referrals by hospital.

Hospital	Referrals
Prince Charles	678
Royal Glamorgan	523
Princess of Wales	226
Morriston	4
Neath Port Talbot	1
Total	**1432**

For this analysis, we focus particularly on Prince Charles Hospital (PCH) in Merthyr Tydfil and Royal Glamorgan Hospital (RGH) in Ynysmaerdy. Princess of Wales Hospital (PoWH) joined the CTMUHB health board in April 2019, resulting in fewer referrals at PCH and RGH.

Table 1.2: frequency table- referral type.

Referral Source	Number of Referrals
Referral from GP	851
Consultant-Internal	266
Emergency	143
Diagnostic-Imaging	56
Other healthcare professional	34
Diagnostic-Other	23
Ward	21
Diagnostic-Endoscopy	15
Outpatients	12
Consultant-External	11
Total	**1432**

Figure 1.6: referrals by day/week/month at PCH.

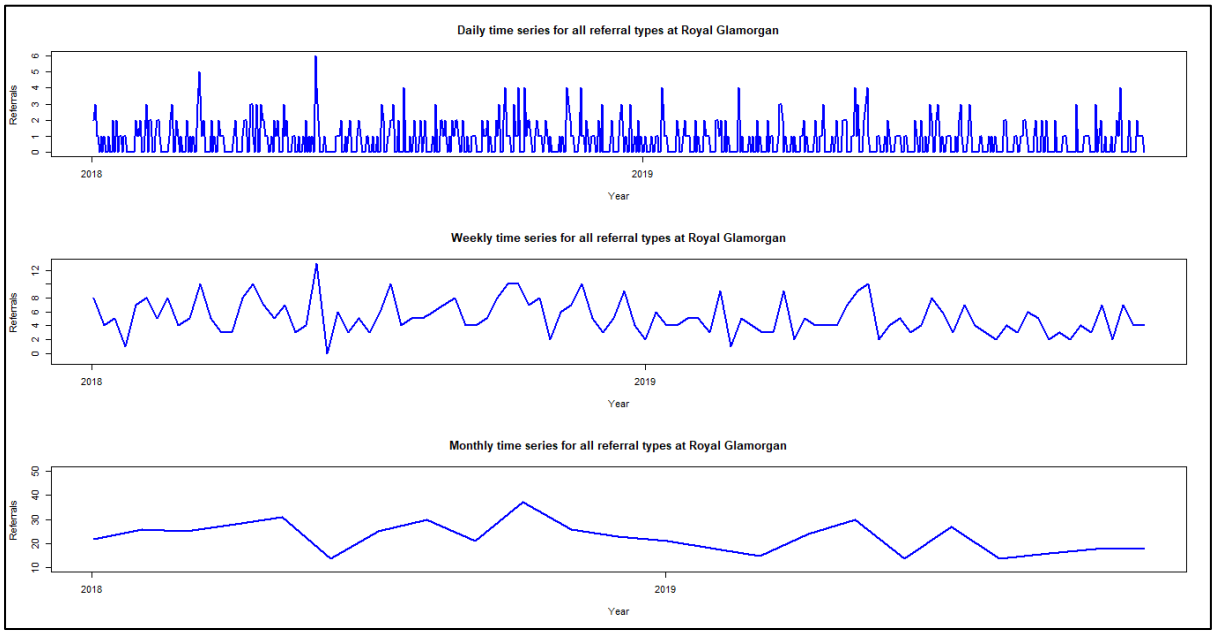

Figure 1.7: referrals by day/week/month at RGH.

From the weekly/monthly data, there are more referrals to PCH than RGH.

From our referral type table, the data is dominated by GP referrals. Many referrals also originate from internal consultants. Other referral types are rare and difficult to forecast

accurately. We aggregate all types of referral, partitioning by hospital. We focus on the *tests* data:

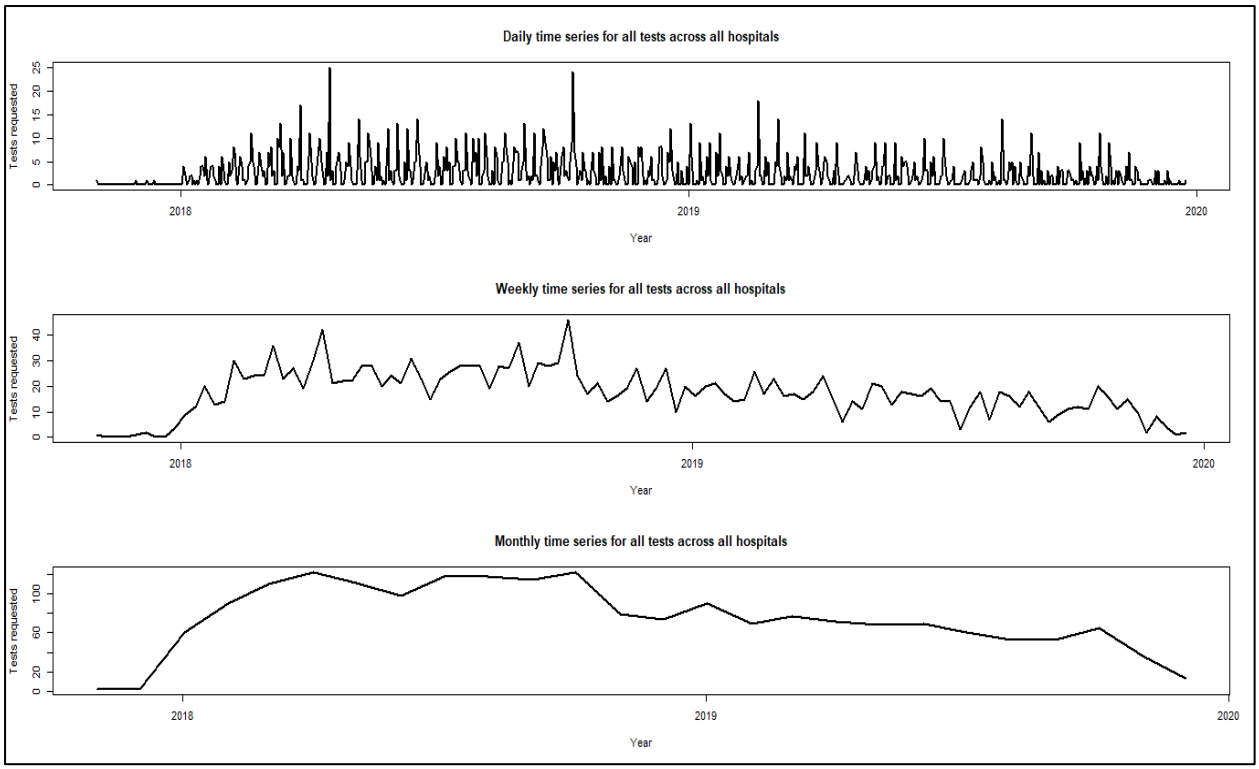

Figure 1.8: daily/weekly/monthly tests requested (all hospitals).

Again, daily data is noisy. There are often zero requests, however sometimes as many as 25. The weekly/monthly data provides a better insight into the structure of the data, which increases until late-2018 before steadily decreasing through 2019. We partition by hospital and test type:

Table 1.3: frequency table- tests by hospital.

Hospital	Test Requests
Prince Charles	1120
Princess of Wales	16
Royal Glamorgan	807
Total	**1943**

Table 1.4: frequency table- test requests by type.

Test Type	Test Requests
CT (excluding Guided Biopsy)	804
Others	382
PET	274
CT Guided Biopsy	203
Bronchoscopy	96
EBUS	74
MRI	74
US Guided Biopsy	36
Total	**1943**

We focus on the time series for CT tests (all hospitals), PET tests (all hospitals), and all other tests combined (all hospitals):

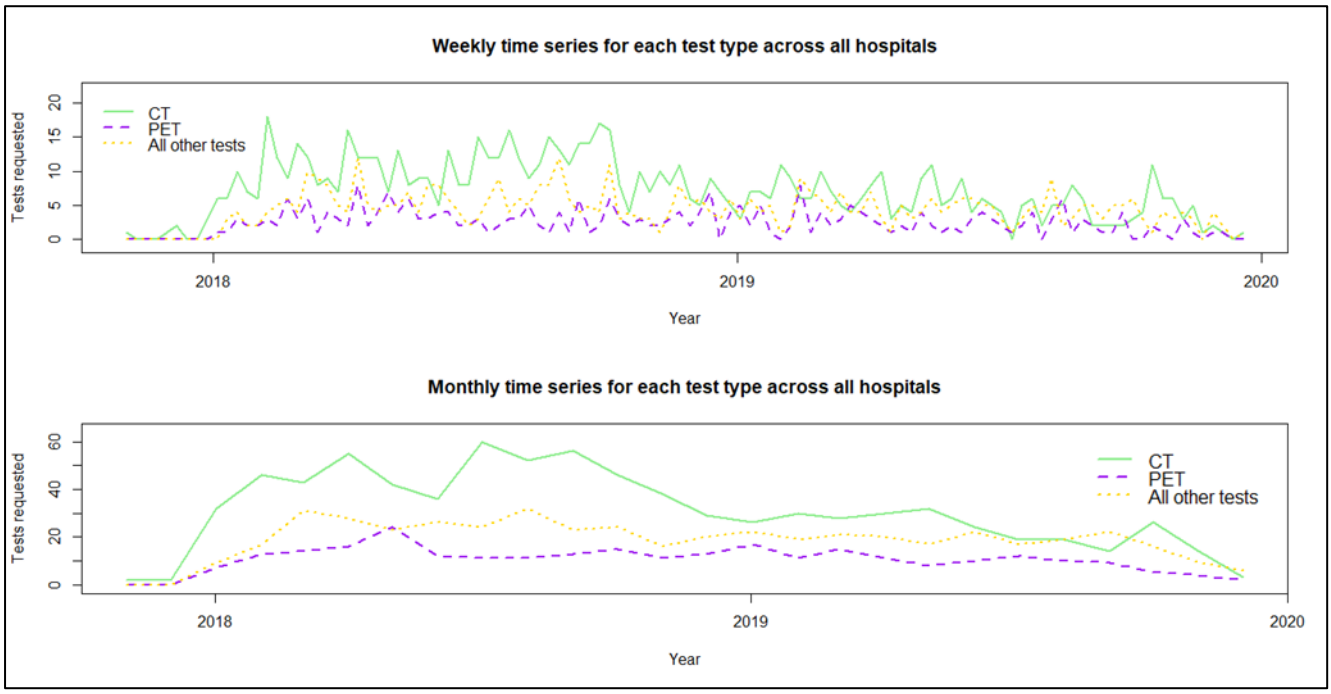

Figure 1.9: weekly/monthly test requests (all hospitals), partitioned by test type.

All the time series we have plotted contain around 2 years of data therefore we only have around 25-30 observations for monthly data. Time series analysis works best with 100+ observations (Box and Tiao 1975)[13]. For higher accuracy forecasting monthly data, the data we have available may be insufficient.

1.3.2 Summary statistics

The *R* command *summary* extracts summary statistics from datasets: minimum, 1st quartile, median, mean, 3rd quartile, and maximum. These hint at the spread of the data before we perform complex analysis. We examine summary statistics for *referrals*:

Table 1.5: summary statistics- daily/weekly/monthly referrals, partitioned by hospital.

Time Series	Min.	1st Quart.	Median	Mean	3rd Quart.	Max.
Daily						
All referrals, all hospitals	0	0	2	2.05	3	11
All referrals, Prince Charles	0	0	1	0.97	2	6
All referrals, Royal Glamorgan	0	0	0	0.75	1	6
Weekly						
All referrals, all hospitals	5	11.75	14	14.32	17	25
All referrals, Prince Charles	1	5	7	6.78	8.25	15
All referrals, Royal Glamorgan	0	3	5	5.23	7	13
Monthly						
All referrals, all hospitals	43	48	61	62.26	70.50	97
All referrals, Prince Charles	16	25.50	29	29.48	33	43
All referrals, Royal Glamorgan	14	18	23	22.74	26.50	37

PCH generally receives more referrals than RGH. However, the two hospitals' busiest days and weeks have seen identical numbers of patients. We analyse the *tests* data:

Table 1.6: summary statistics- daily/weekly/monthly tests, partitioned by hospital.

Time Series	Min.	1st Quart.	Median	Mean	3rd Quart.	Max.
Daily						
All tests, all hospitals	0	0	1	2.48	4	25
Weekly						
All referrals, all hospitals	0	12	17.50	17.35	23	46
CT tests, all hospitals	0	4	7	7.18	10	18
PET tests, all hospitals	0	1	2	2.45	4	8
Other tests, all hospitals	0	3	4	4.31	6	12
Monthly						
All referrals, all hospitals	2	60.25	72.50	74.73	107	122
CT tests, all hospitals	2	20.25	30	30.92	42.75	60
PET tests, all hospitals	0	8.25	11	10.54	13	24
Other tests, all hospitals	0	16.25	20	18.58	23	32

CT tests are the most frequent tests being requested.

1.3.3 Boxplots

We create boxplots to visualise the data according to its value for day/month. Such plots allow us to assess which days/months are typically busy/idle:

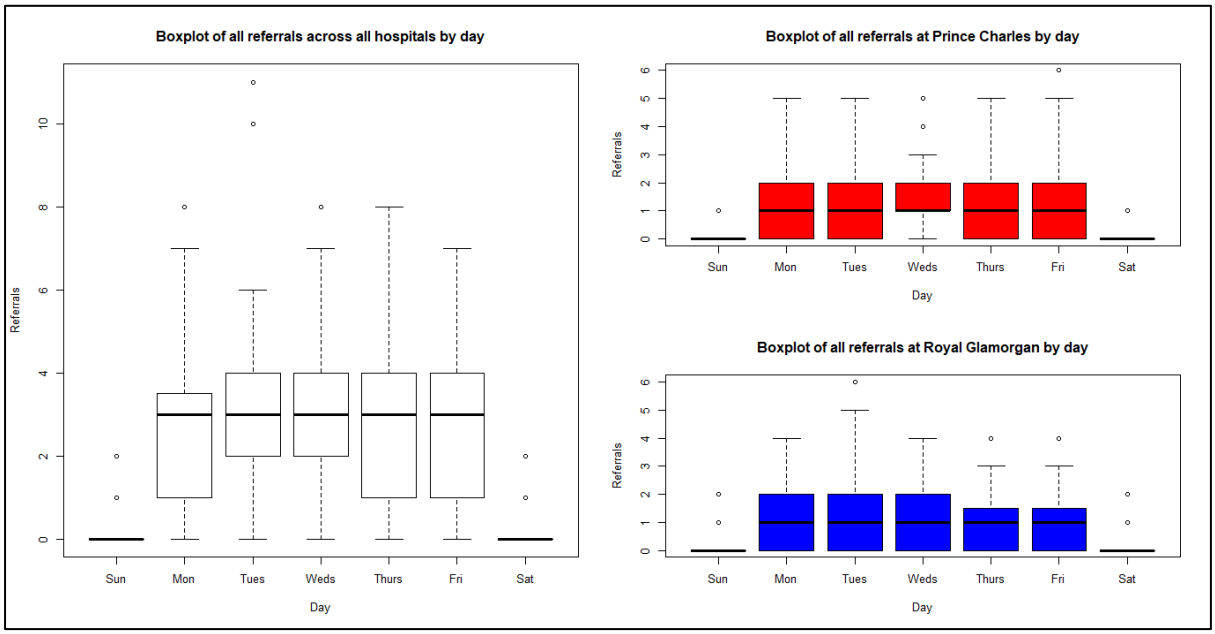

Figure 1.10: boxplots- referrals by day (all hospitals) and at PCH and RGH.

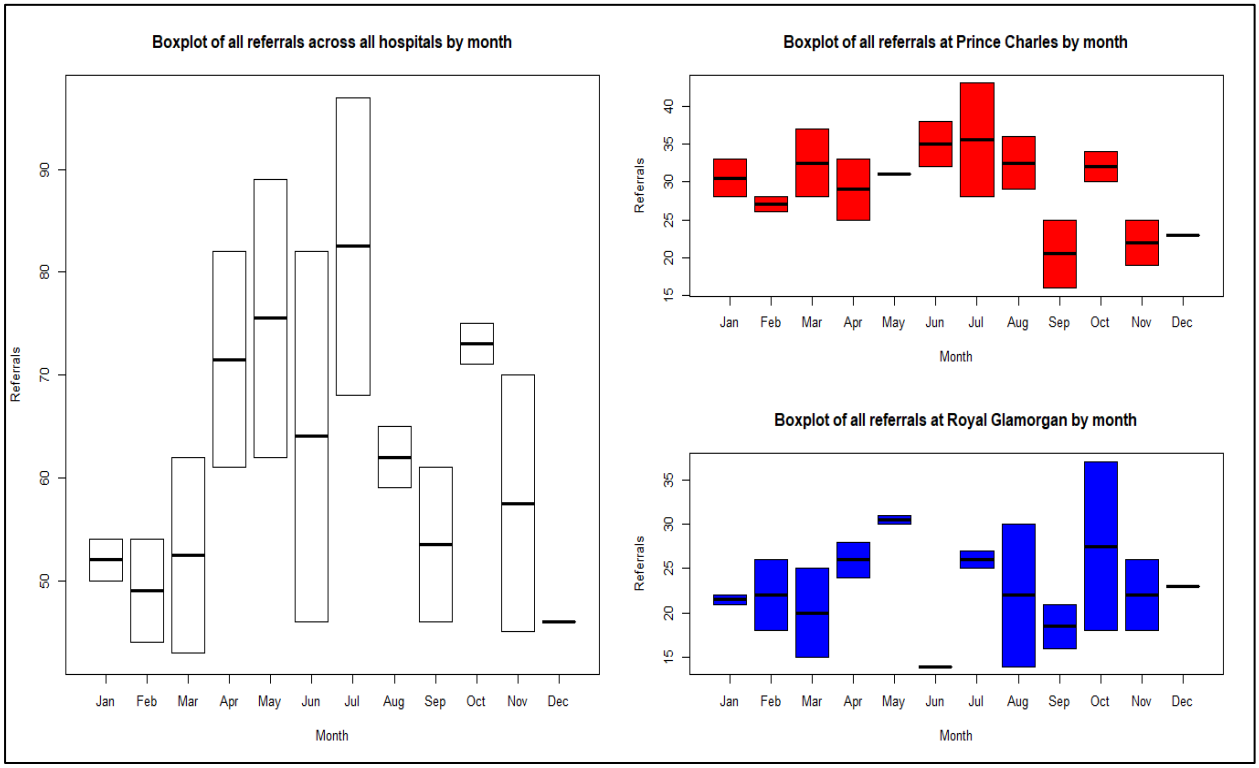

Figure 1.11: boxplots- referrals by month (all hospitals) and at PCH and RGH.

We identify consistent referral numbers throughout the working week; weekend activity is rare. We also see more activity in spring/summer. Analysis of monthly data is difficult, due to its scarcity (our time series is under 2 years). We create boxplots for the *tests* data, focussing on CT, PET and 'other' test types.

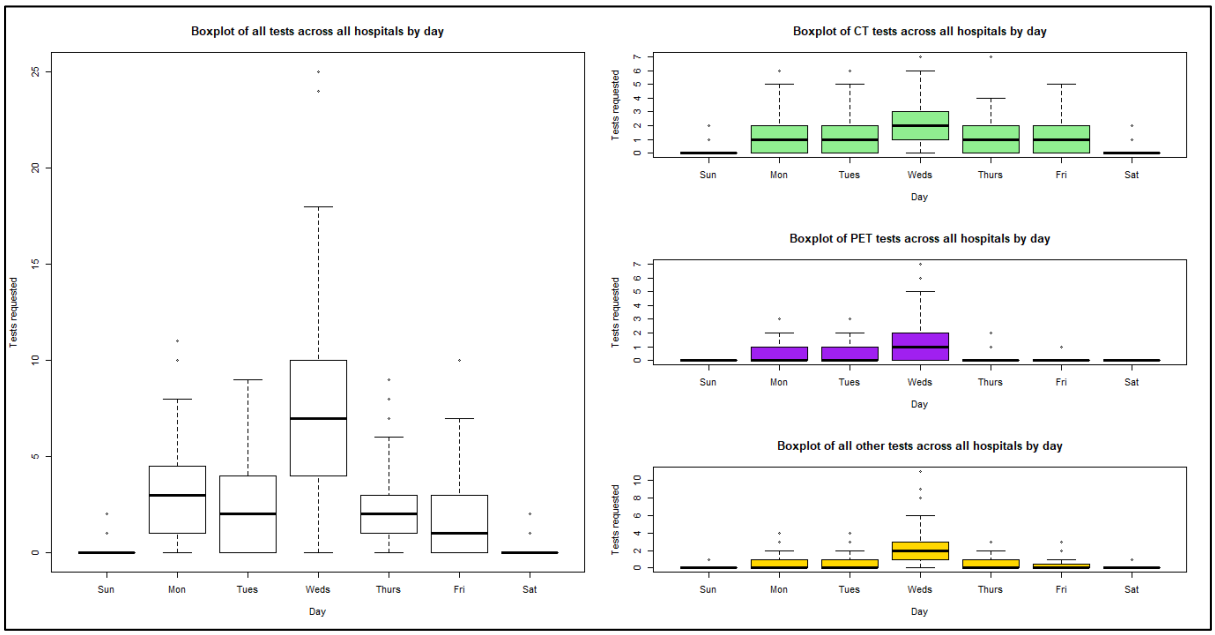

Figure 1.12: boxplots- tests by day (all hospitals).

For all days excluding Wednesday, we have mean zero PET or 'other' test requests.

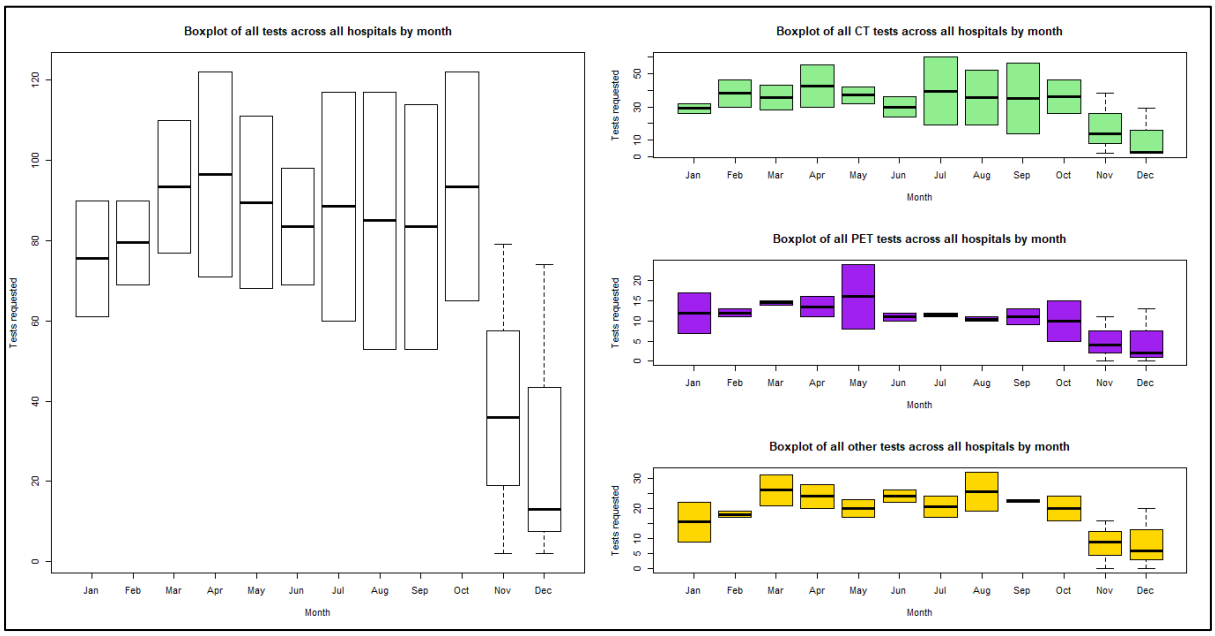

Figure 1.13: boxplots- tests by month (all hospitals).

Again, weekend activity is uncommon. This is unsurprising, many medical professionals only work weekdays. The number of tests decreases towards the year-end. These boxplots should be used for further investigation into seasonal variation.

1.3.4 Moving averages (MAs)

MAs calculate the mean of a fixed subset and then shift forward, taking the mean of subsequent subsets. These calculations smooth fluctuations. For a time series $Y_1, Y_2, \ldots, Y_n$ we define the MA of period k (MA(k)) as:

$$\frac{(Y_1 + Y_2 + \cdots + Y_k)}{k}, \frac{(Y_2 + \cdots + Y_{k+1})}{k}, \frac{(Y_3 + \cdots + Y_{k+2})}{k}, \ldots$$

[14]MAs smooth seasonality from weekly/monthly variation. We use the *R* command *ma*, available through the *forecast* package available in *R*. We plot our MAs by 7 days for our daily data, 4 weeks for our weekly data and 3 months for our monthly data. These plots show all referrals/tests, with moving averages (including partitioned data).

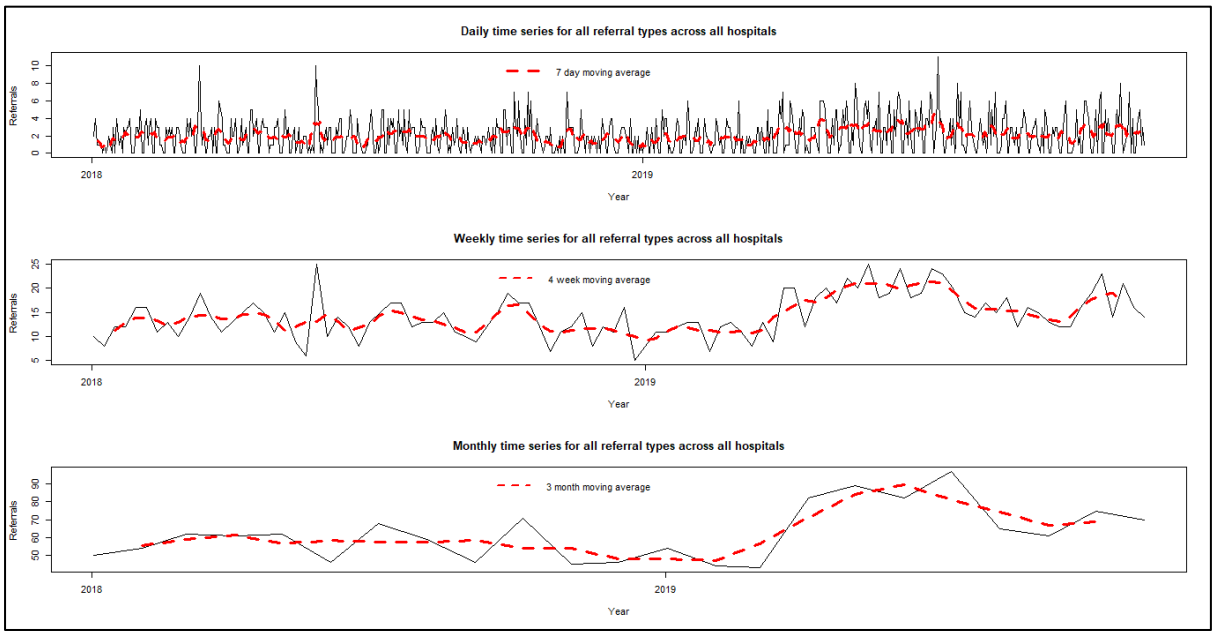

Figure 1.14: moving averages- daily/weekly/monthly referrals (all hospitals).

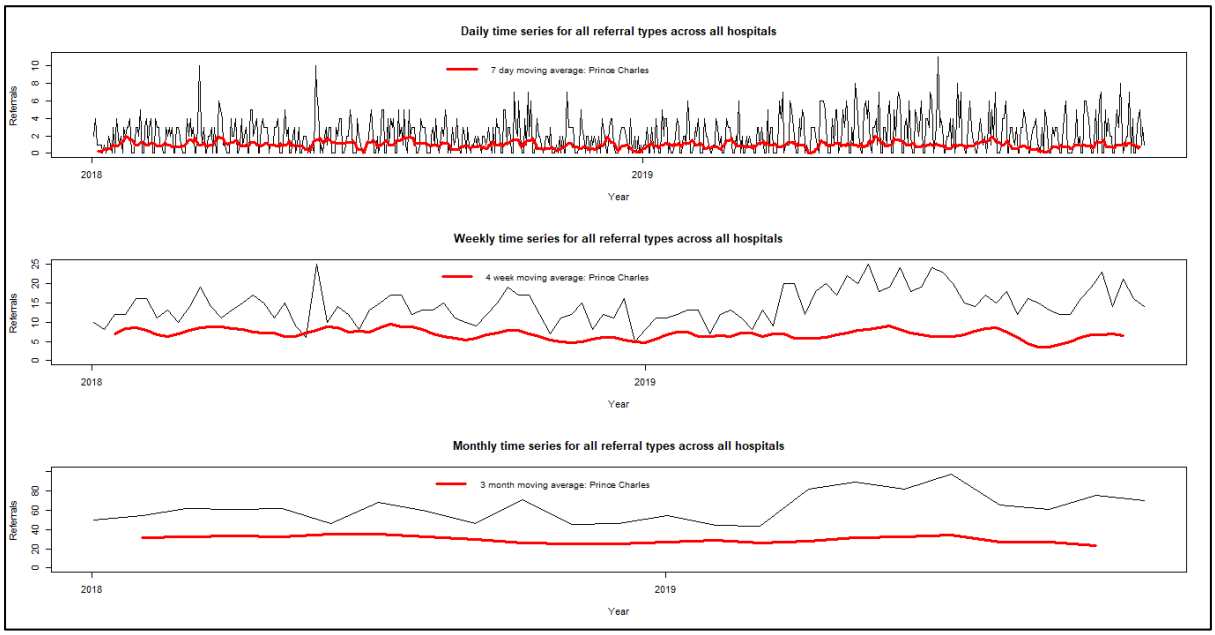

Figure 1.15: moving averages- daily/weekly/monthly referrals (all hospitals).

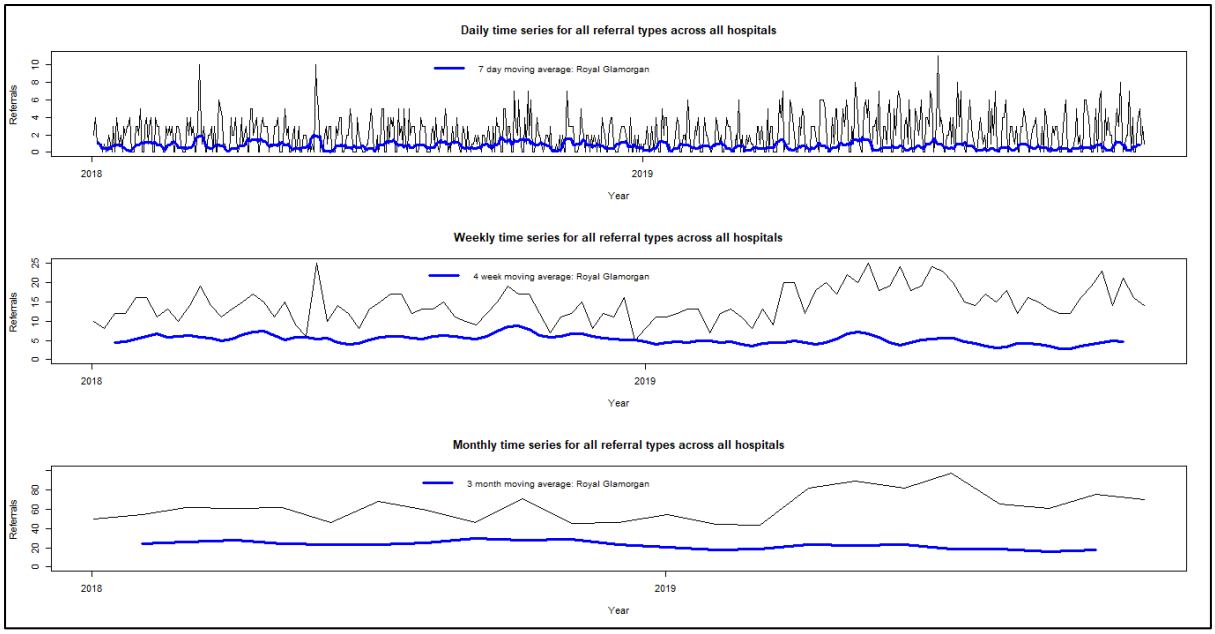

Figure 1.16: moving averages- daily/weekly/monthly referrals.

We next consider the *tests* data:

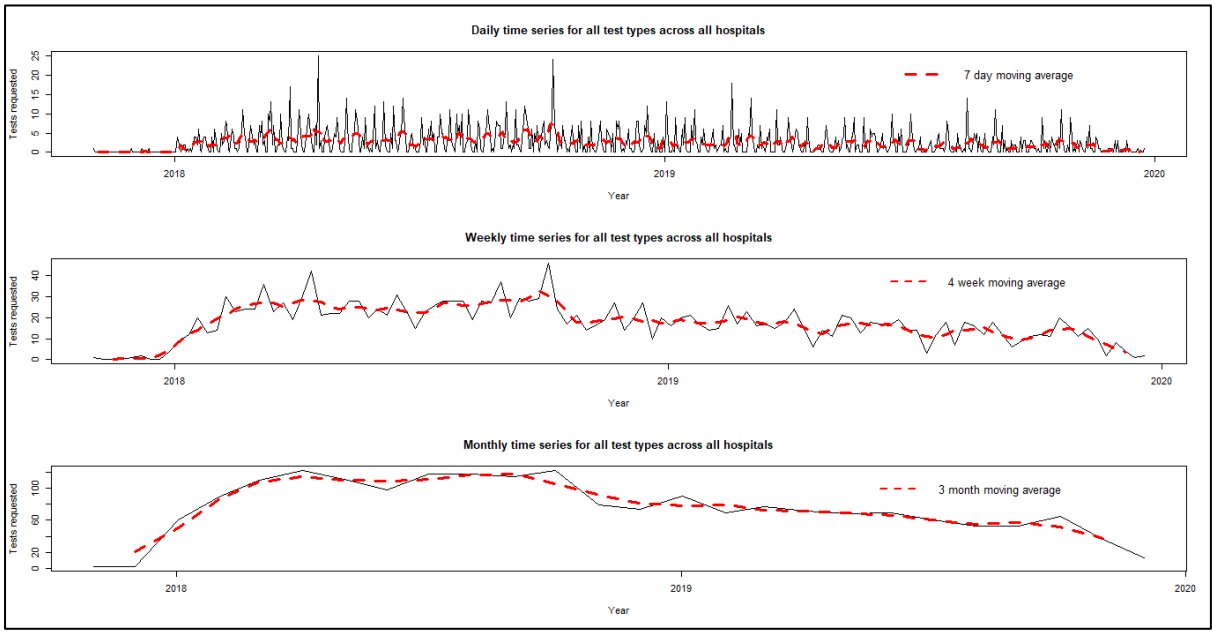

Figure 1.17: moving averages- daily/weekly/monthly tests (all hospitals).

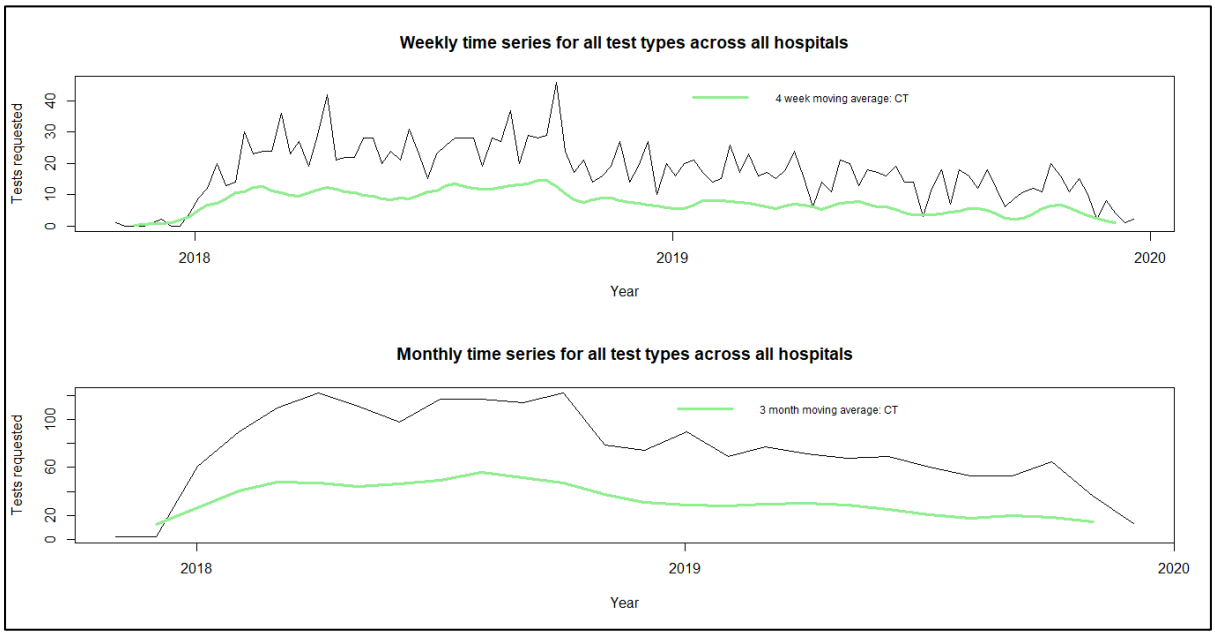

Figure 1.18: moving averages- weekly/monthly tests (all hospitals).

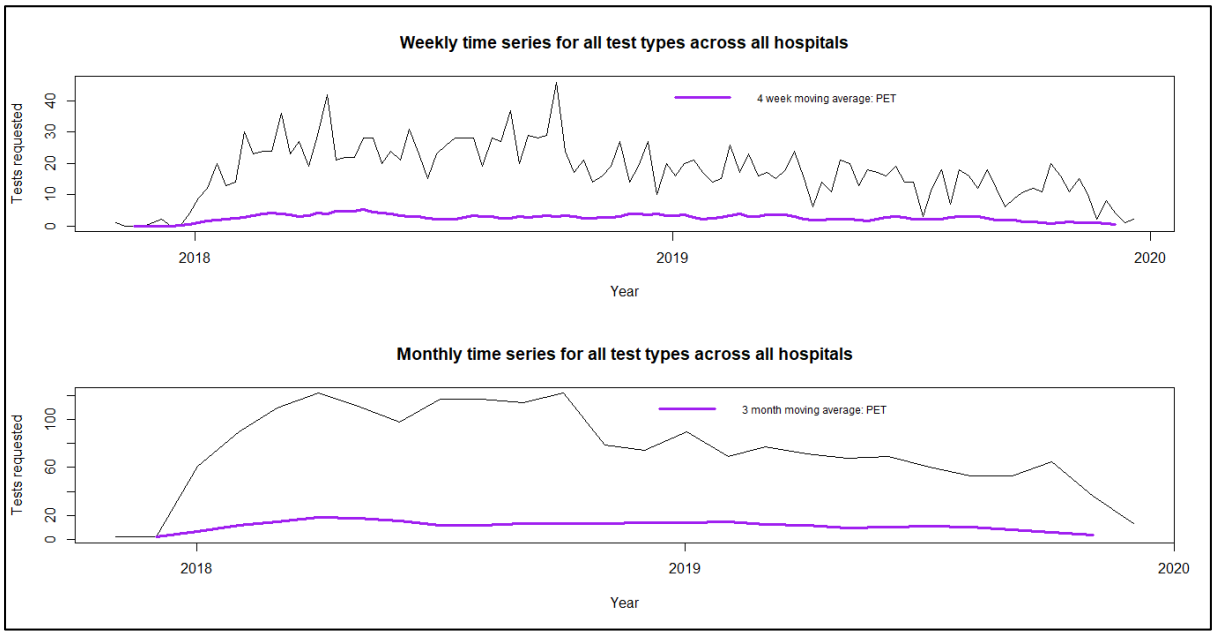

Figure 1.19: moving averages- weekly/monthly tests (all hospitals).

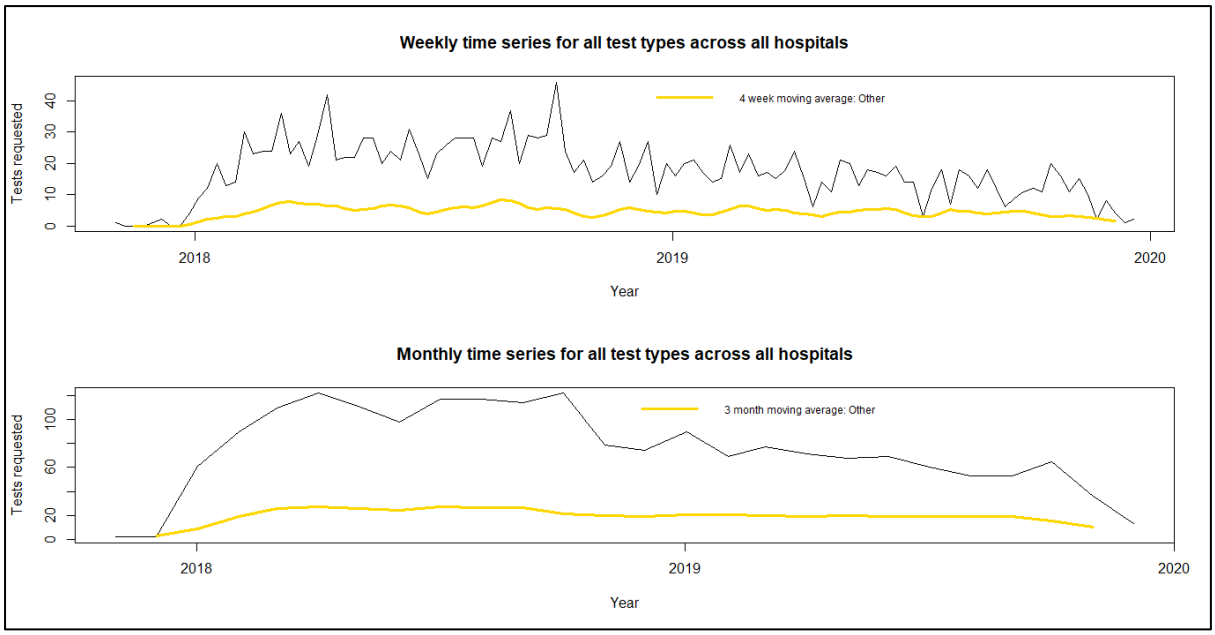

Figure 1.20: moving averages- weekly/monthly tests (all hospitals).

Tests increase through early 2018 and plateau until late-2018. Afterwards, test numbers steadily decrease. Again, this data should be used for further investigation.

1.3.5 Stationarity

A stationary time series has a constant mean, variance, and autocorrelation. For a stationary series, forecasted values reflect previous observations. To determine stationarity, we require the *adf.test* function in *R*, this function performs an *augmented Dickey-Fuller* (ADF) *Test* for stationarity and is available through the *tseries* package. The ADF test searches for a unit root

in the time series, implying nonstationarity. We test the null hypothesis of nonstationarity against the alternative hypothesis of stationarity. Our daily data contains many zeroes which skew results; therefore, we examine weekly/monthly data for referrals/tests.

We convert our data frames into a *ts* object in *R* using the *ts* command, which takes arguments including the dataframe, start date, end date, and frequency (observations per unit of time)[15]. For daily data, we use frequency 365 (365 days in a year). For weekly data, we use frequency 52 (52 weeks per year). For monthly data, we use frequency 12 (12 months per year).

The *referrals* data is slightly short of 2 years, therefore we set frequency to 349 days (daily), 50 weeks (weekly) and 11 months (monthly). These figures represent the lowest value for 2 periods. This was necessary as many time series operations in *R* require at least 2 periods of data. This adjustment, however, impacts the accuracy of forecasts and could be rectified by collecting more data. The *tests* data was unaffected by this issue and therefore had frequency 365 days, 52 weeks, and 12 months.

The ADF test statistic is negative. The statistic's magnitude correlates with the confidence of stationarity for our series[16]. In this analysis, we use a confidence level of 95%, suggesting that p-values above 0.05 are statistically significant. We perform the test on the *referrals* data*:*

Table 1.7: augmented Dickey-Fuller tests on referrals.

Time series	ADF test statistic	Lag order	p-value	Suspected stationarity
Daily				
All referrals, all hospitals	-7.45	8	0.01	Stationary
Weekly				
All referrals, all hospitals	-2.59	4	0.33	Nonstationary
All referrals, Prince Charles	-4.84	4	0.01	Stationary
All referrals, Royal Glamorgan	-5.19	4	0.01	Stationary
Monthly				
All referrals, all hospitals	-1.83	2	0.64	Nonstationary
All referrals, Prince Charles	-2.19	2	0.50	Nonstationary
All referrals, Royal Glamorgan	-2.59	2	0.35	Nonstationary

We deduce that the weekly/monthly referrals data for all hospitals is nonstationary. Interestingly the weekly referrals series at PCH and RGH scored large test statistics and low p-values, suggesting stationarity. The weekly series for PCH and RGH appears stationary. To examine long-term behaviour, we consider the augmented Dickey-Fuller test results on the

monthly series. The monthly data for all referrals is nonstationary whether partitioned by hospital or not, implying that the series are long-term nonstationary. We examine the tests data:

Table 1.8: augmented Dickey-Fuller tests on tests.

Time series	ADF test statistic	Lag order	p-value	Suspected stationarity
Daily				
All tests, all hospitals	-6.07	9	0.01	Stationary
Weekly				
All tests, all hospitals	-2.81	4	0.24	Nonstationary
CT tests, all hospitals	-3.17	4	0.10	Nonstationary
PET tests, all hospitals	-2.93	4	0.19	Nonstationary
Other tests, all hospitals	-2.79	4	0.25	Nonstationary
Monthly				
All tests, all hospitals	-3.73	2	0.04	Stationary
CT tests, all hospitals	-3.46	2	0.07	Nonstationary
PET tests, all hospitals	-3.23	2	0.10	Nonstationary
Other tests, all hospitals	-3.80	2	0.04	Stationary

We again suspect stationarity of the daily data. Interestingly, we fail to reject the null hypothesis of nonstationarity for all tests (weekly). We, however, suspect stationarity for all tests (monthly). Our analysis suggests long-term nonstationarity for CT and PET tests.

1.3.6 Decomposition: trend/seasonality/error

We examine the daily/weekly/monthly decompositions all referrals (all hospitals). Restricting our interest to all *referrals* and all *tests*:

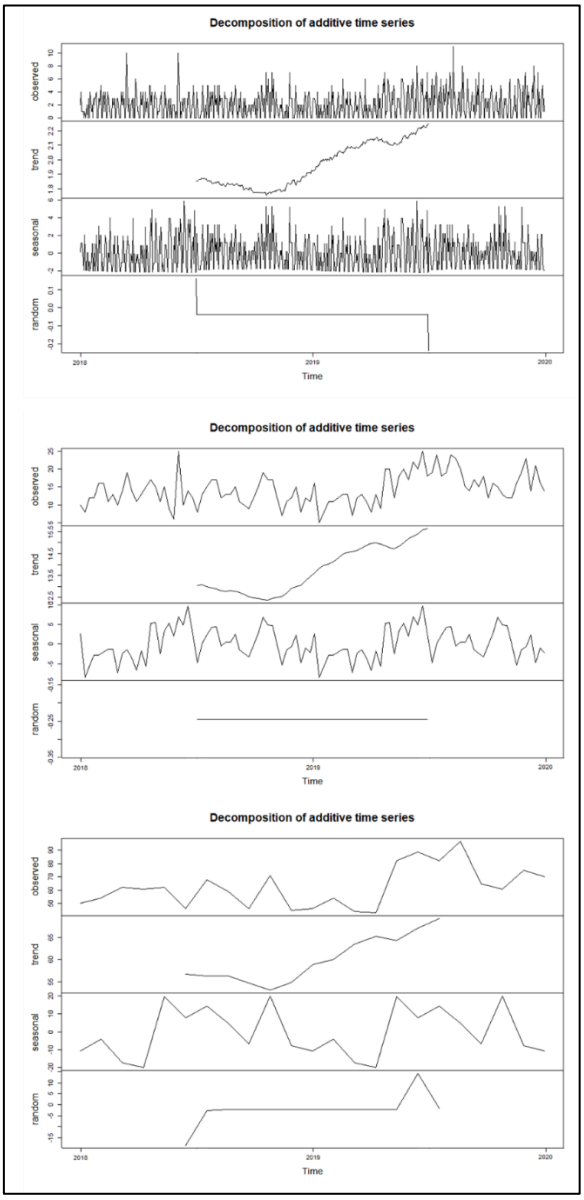

Figure 1.21: decomposition- daily/weekly/monthly time series for all referrals (all hospitals).

Again, measuring over a longer period lessens noise. In all three plots, we see an increasing trend consistent with our earlier finding of nonstationarity in the weekly/monthly data for all referrals (all hospitals). The seasonality plot contains two periods, matching the periods in our series. Examine decomposition of *referrals* by hospital:

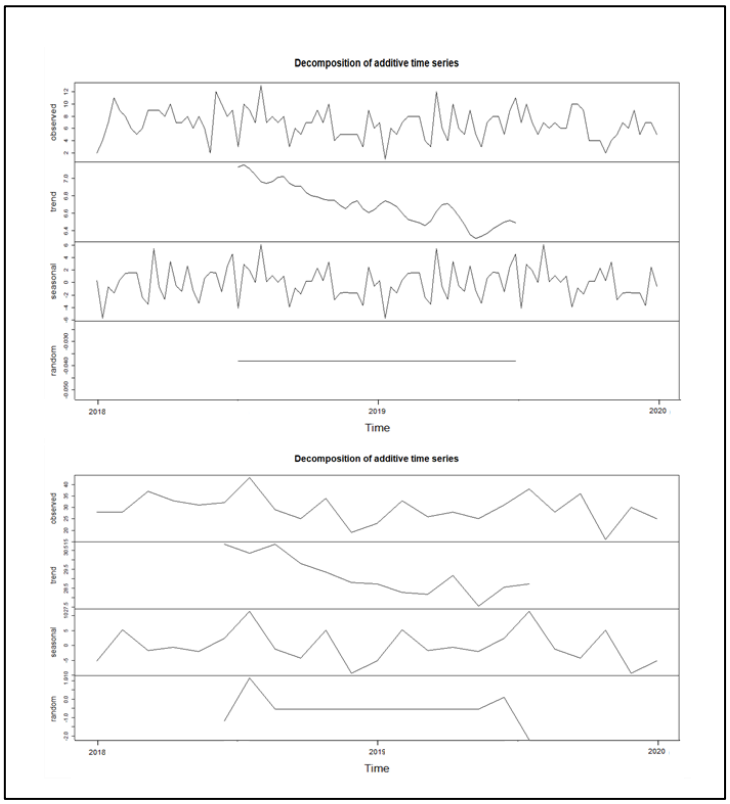

Figure 1.22: decomposition- weekly/monthly time series for all referrals (PCH).

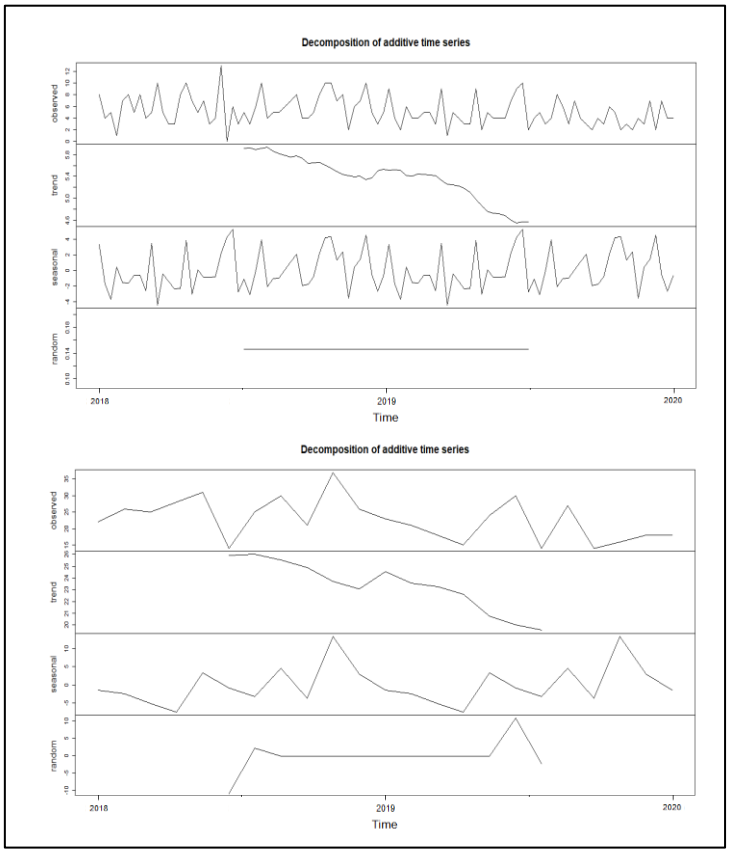

Figure 1.23: decomposition- weekly/monthly time series for all referrals (RGH).

The trends for all hospitals are decreasing. We next decompose the weekly series for the 3rd busiest hospital: PoWH.

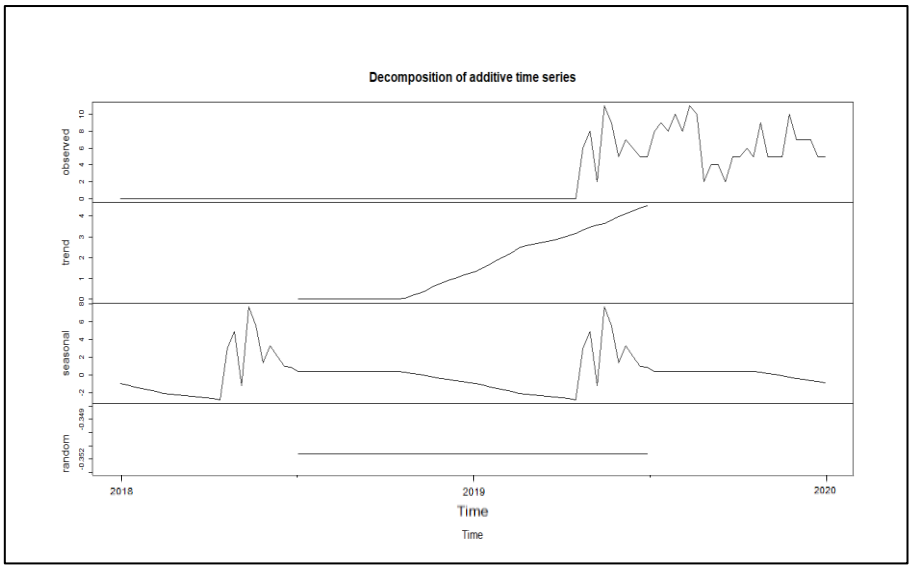

Figure 1.24: decomposition- PoWH weekly series.

After the PoWH begins receiving referrals in early-2019, referrals rapidly increase. This suggests that, from spring-2019, referrals to PCH or RGH were directed instead to PoWH. We examine the *tests* data:

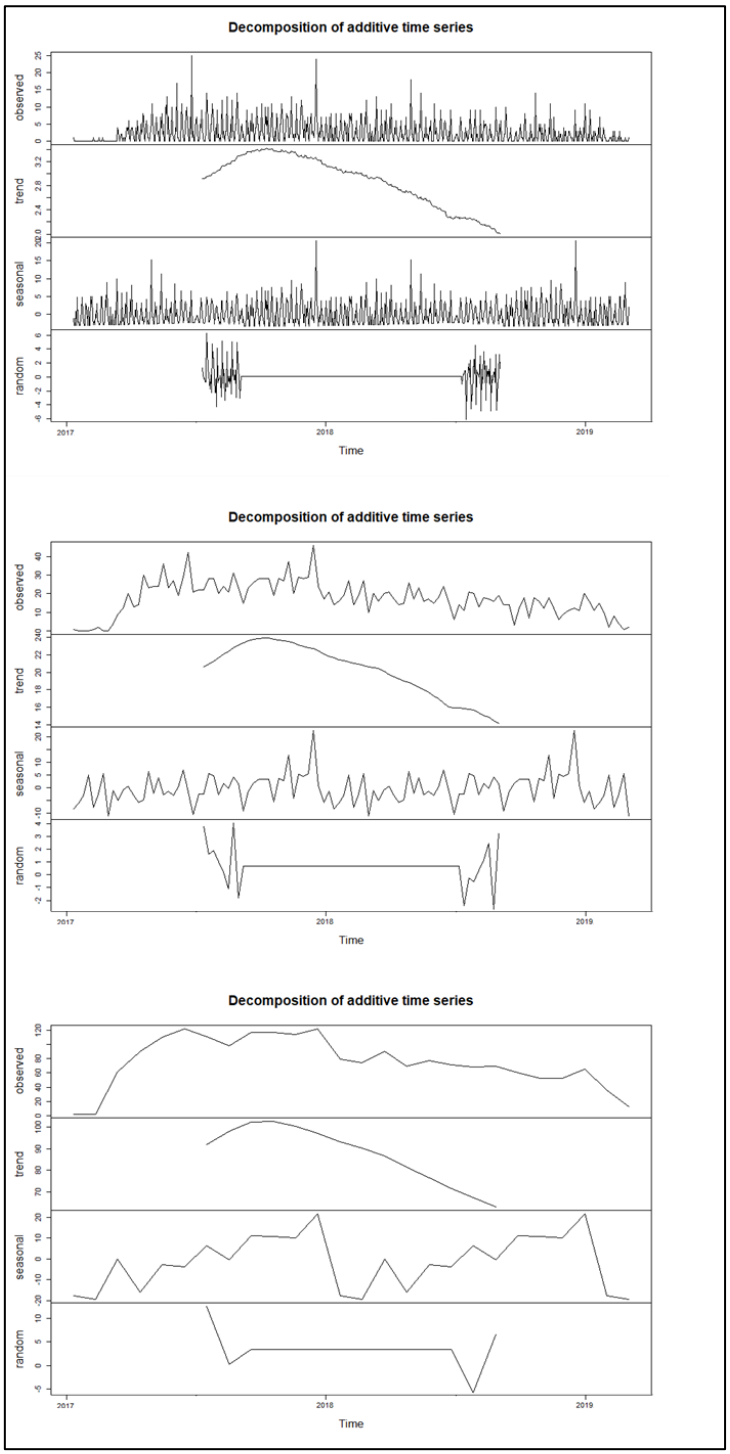

Figure 1.25: decomposition- daily/weekly/monthly time series for all tests (all hospitals).

The weekly/monthly data has seasonality from two periods. The trend appears to decrease after an initial increase. Our stationarity tests suggest no statistical evidence for stationarity on the weekly data but suggest nonstationarity on the monthly data. Therefore, the tests are decreasing long-term. We should consult with the health board to investigate this decrease.

The shortness of our data impacts on the random error element of the decomposition. We expect randomly distributed random errors with mean zero, however our error diagnostic stays constant. This was rectified by decreasing the frequency of the time series, suggesting that more data would result in a better-behaved error component. We next discuss different forecasting methods for our series.

1.4 Discussion of methods

We discuss simple forecasting methods to predict future observations, beginning by focussing on the *simple naïve* model. We then examine more complex methods including extrapolation models, regression models, ARIMAs, SSA and ANNs. For brevity, we focus on the series for all referrals to all hospitals, and all tests of all types. We shall, however, apply the methods to all our partitioned datasets and consider the findings.

The *accuracy* function, available through the forecast package in *R*, calculates statistical information regarding the residuals (error) in our models.

We partition our series into a training set (75%), and test set (25%). The training set builds the model and the test set measures its effectiveness. Such a split allows adequate training for our models but reserves a large enough portion for valid accuracy tests. We begin with the simple naïve forecast method, often used as a baseline.

1.4.1 Simple naïve model

$\hat{Y}_{t+1} = Y_t$

The naïve model assumes the forecast for time $t + 1$ is equal to the previous observation. We fit a naïve model using the *naïve* function from the *forecast* package in *R*. We create several

plots recording the forecast statistics for all our series. *R* plots the range of the 80% and 95% confidence levels for each forecast.

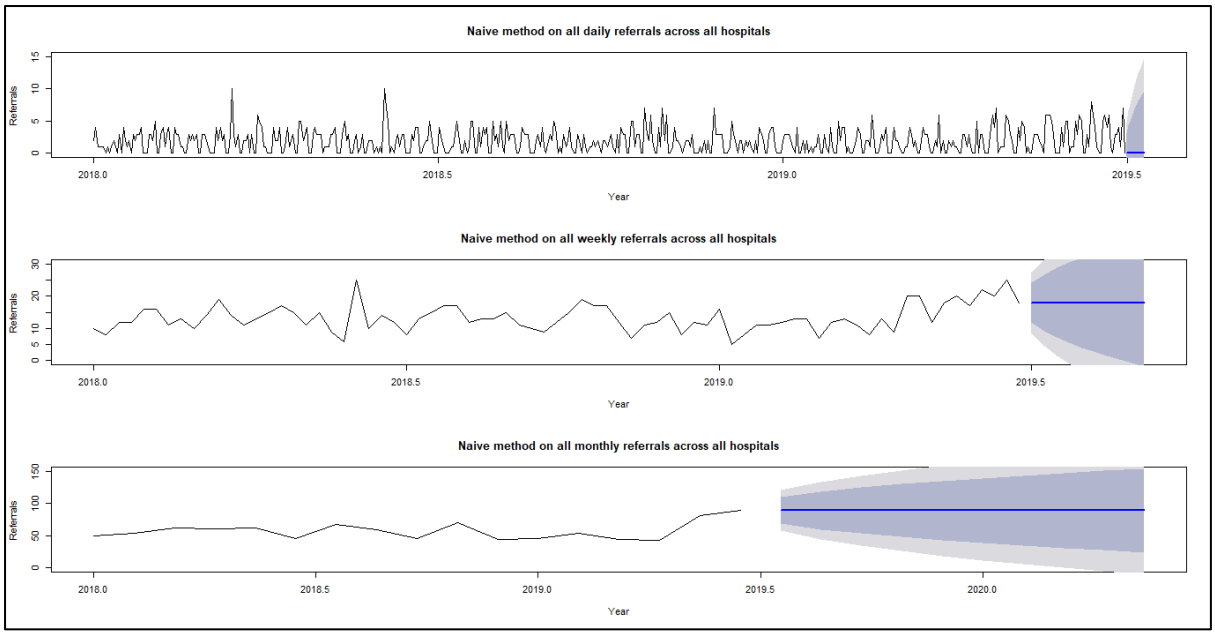

Figure 1.26: naïve model- daily/weekly/monthly referrals (all hospitals).

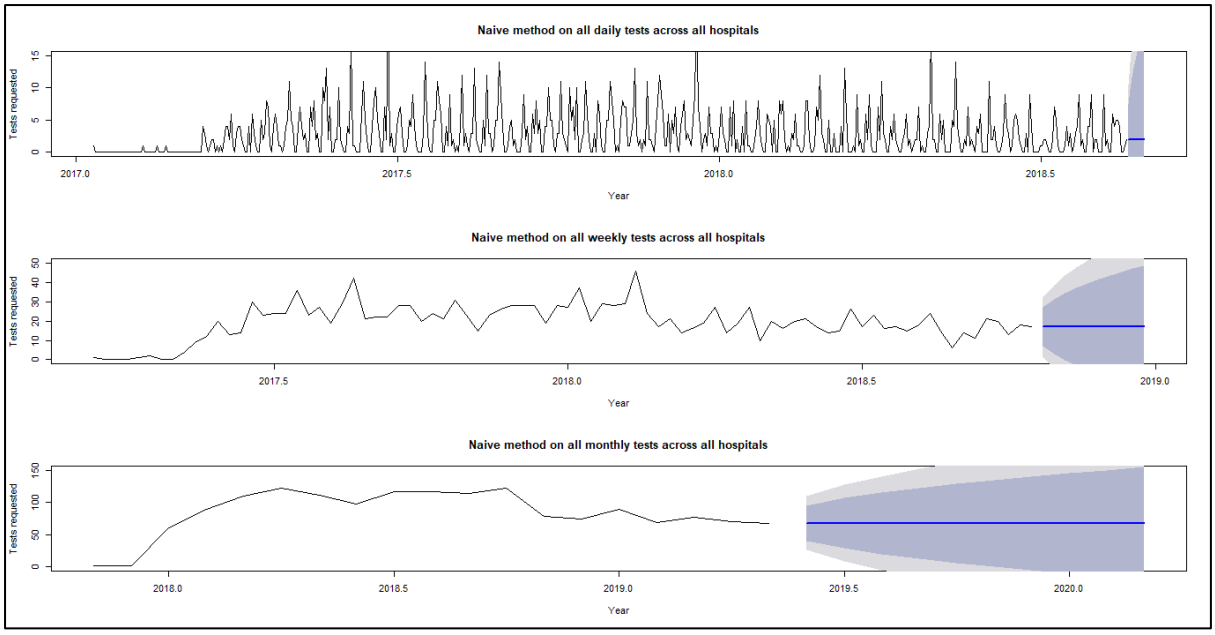

Figure 1.27: naïve model- daily/weekly/monthly tests (all hospitals).

The naïve model's strength is its simplicity; however, it performs poorly on series with increasing/decreasing trends. Despite its crudeness, the naïve method often outperforms more sophisticated methods. Over 43 periods of semi-annual U.S. Treasury bond yields, forecasts of change of yield direction have been incorrect around 65% of the time. Naïve forecasting

reduced the standard deviation in forecast error by 11% Brooks and Gray (2004)[17]. While usually used as a baseline to assess other models, the naïve model may itself be favoured.

The simple naïve model ignores seasonality; therefore, it is unsuitable for many series. We thus examine the *seasonal naïve* model.

1.4.2 Seasonal naïve model

$\hat{Y}_{t+1} = Y_{t-k}$ for seasonal lag k

This model improves the simple naïve model, including a seasonal component. Forecasts are generated from previous observations from the same season. We examine forecasts from the seasonal naïve model on our *referrals* series (all hospitals). We use the *snaive* function:

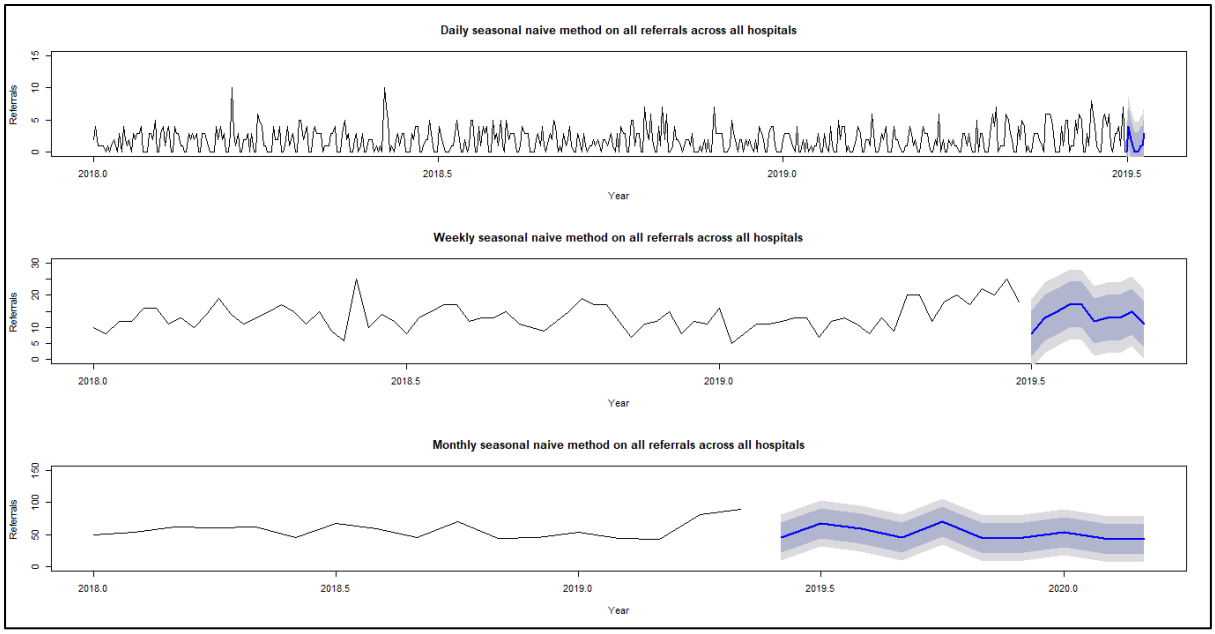

Figure 1.28: seasonal naïve model- daily/weekly/monthly referrals (all hospitals).

We consider *tests*:

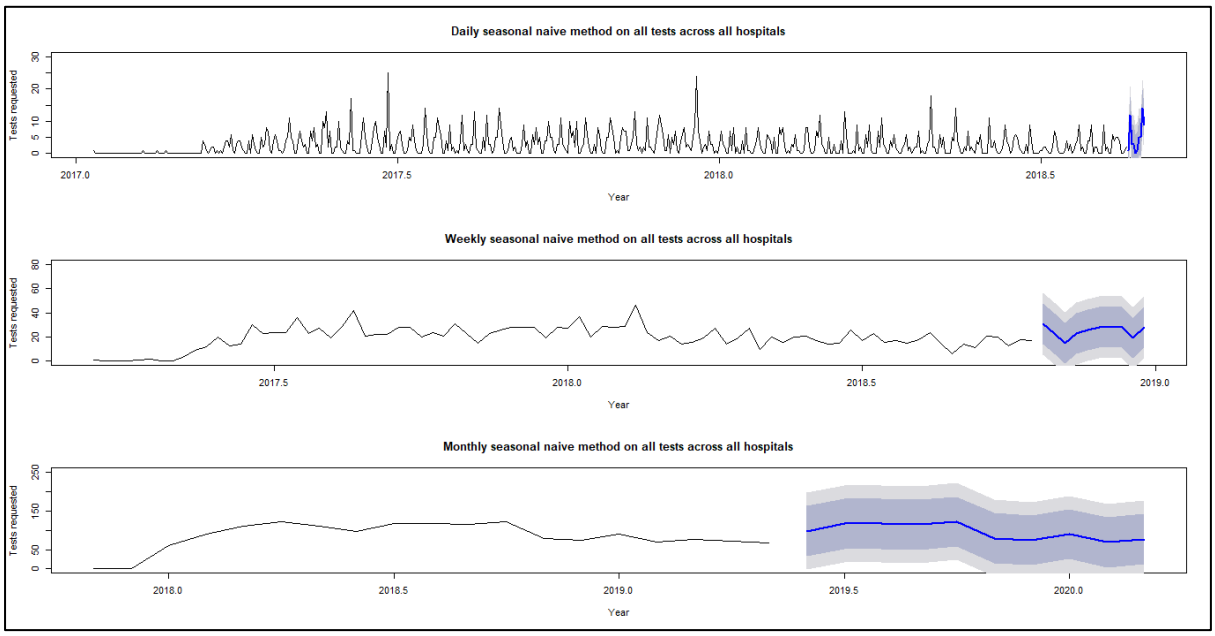

Figure 1.29: seasonal naïve model- daily/weekly/monthly tests (all hospitals).

The seasonal naïve approach offers an improvement on the simple naïve method, factoring seasonality. Again, the method is ineffective for data with a global increasing or decreasing trend, including series for which we already suspect nonstationarity.

We next examine more complex *extrapolation models*.

1.4.3 Extrapolation models

We consider three models: *single exponential smoothing* (SES), *Holt-linear* and *Holt-Winters*. The general form for an extrapolation model is:

$$\hat{Y}_{t+1} = f(Y_t, Y_{t-1}, Y_{t-2}, \dots)$$

Forecasts are produced from some function on preceding observations. f depends on which method is chosen[18].

Single exponential smoothing

First proposed by Robert Goodell Brown in his 1956 paper: *exponential smoothing for predicting demand*, this approach assumes that the series fluctuates around a constant level, or changes level slowly[19]. SES forecasts are given by:

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t)$$

Our forecast, $\hat{Y}_{t+1}$, results from the previous forecast plus some smoothing parameter for level α, $0 \leq \alpha \leq 1$, based on the error from the previous forecast. The choice of smoothing parameters affects the optimisation of the model. One selects a smoothing parameter which minimises either the *sum of squared one-step-ahead forecast errors* (SSE) or *sum of absolute one-step-ahead forecast errors* (SAE). The SAE metric should be used in data with many outliers[20]. We use the *ses* function to fit this model and allow *R* to calculate the smoothing parameter based on SSE.

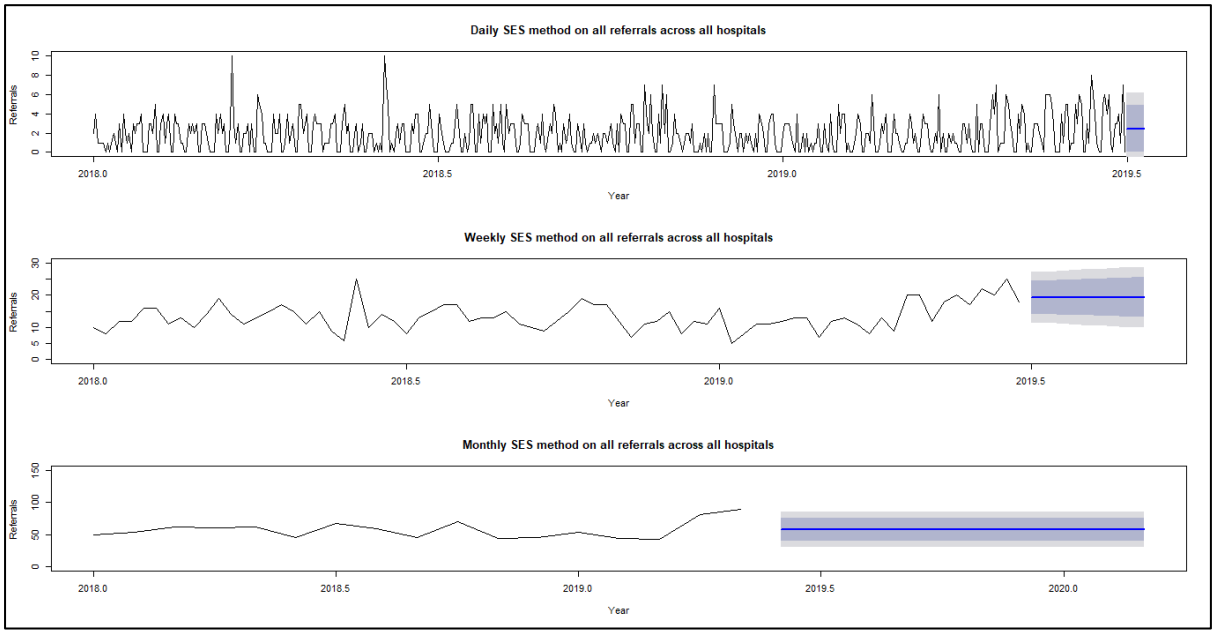

Figure 1.30: SES- daily/weekly/monthly referrals (all hospitals).

As with the naïve method, this method forecasts a constant number of future referrals, usually approximately the series' mean.

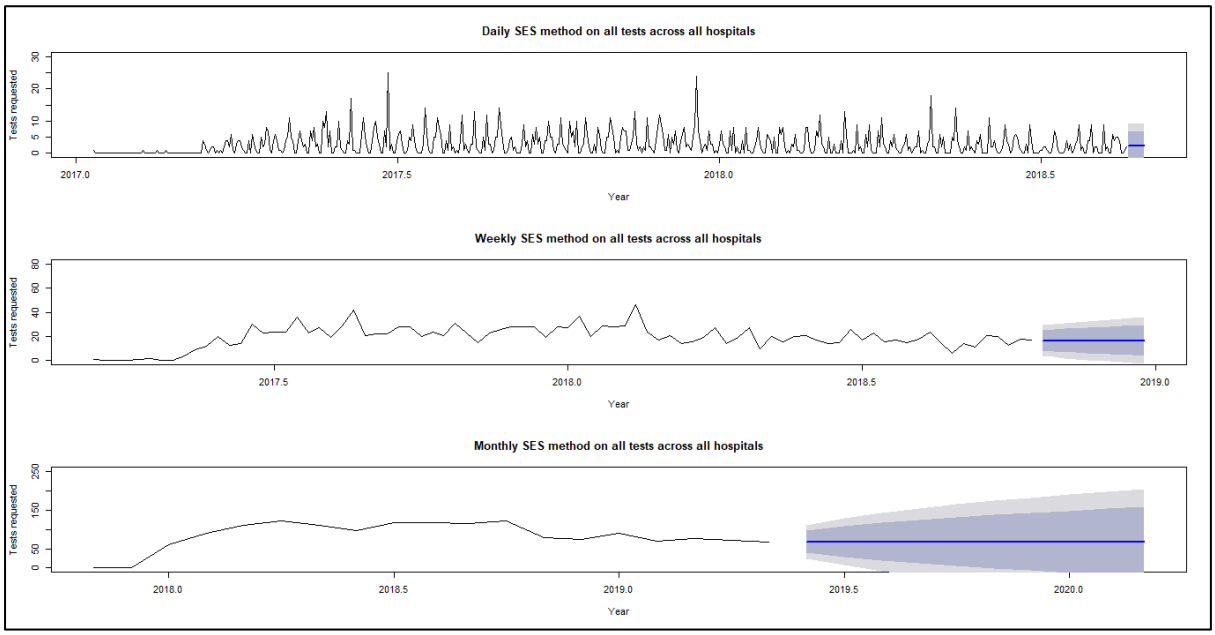

Figure 1.31: SES- daily/weekly/monthly tests (all hospitals).

Exponential smoothing performs well for short-medium forecasts for stationary data[21]. However, it relies on an accurate α, which is difficult to calculate for large data. We also become less confident in the long-term accuracy of any forecast due to uncertainty whether historical patterns will continue indefinitely.

Holt-linear method

Proposed by Charles Holt in 1957[22], this method builds on the SES method, accounting for data with a linear trend. Using the value of the time series at time t, we estimate the base level of the series (E_t) and trend per time period (T_t). Holt's linear method is calculated by the function:

$$\hat{Y}_{t+n} = E_t + nT_t$$

Where

$$E_t = \alpha Y_t + (1-\alpha)(E_{t-1} + T_{t-1})$$
$$T_t = \beta(E_t - E_{t-1}) + (1-\beta)T_{t-1}$$

This provides forecasts for $n \in \mathbb{N}$ future periods. Again, smoothing parameters are calculated to minimise the SSE. The crucial difference between SES and Holt-linear is the

trend function, including smoothing parameter: β, $0 \leq \beta \leq 1$. Applying this method to our *referrals*/*tests* data:

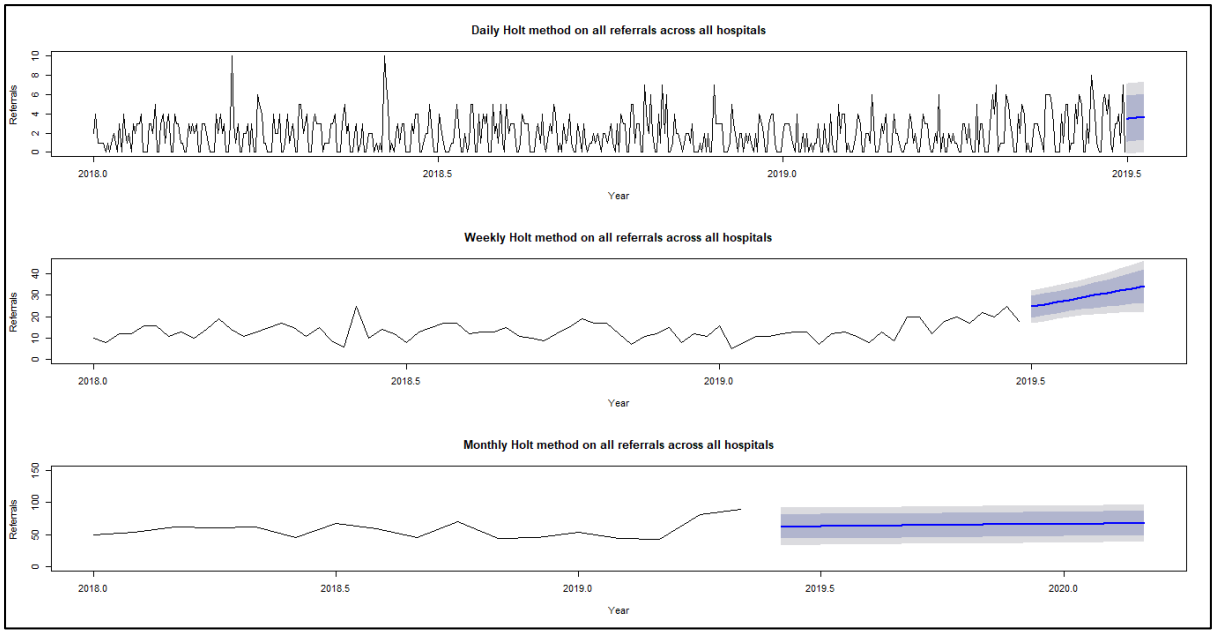

Figure 1.32: Holt-linear- daily/weekly/monthly referrals (all hospitals).

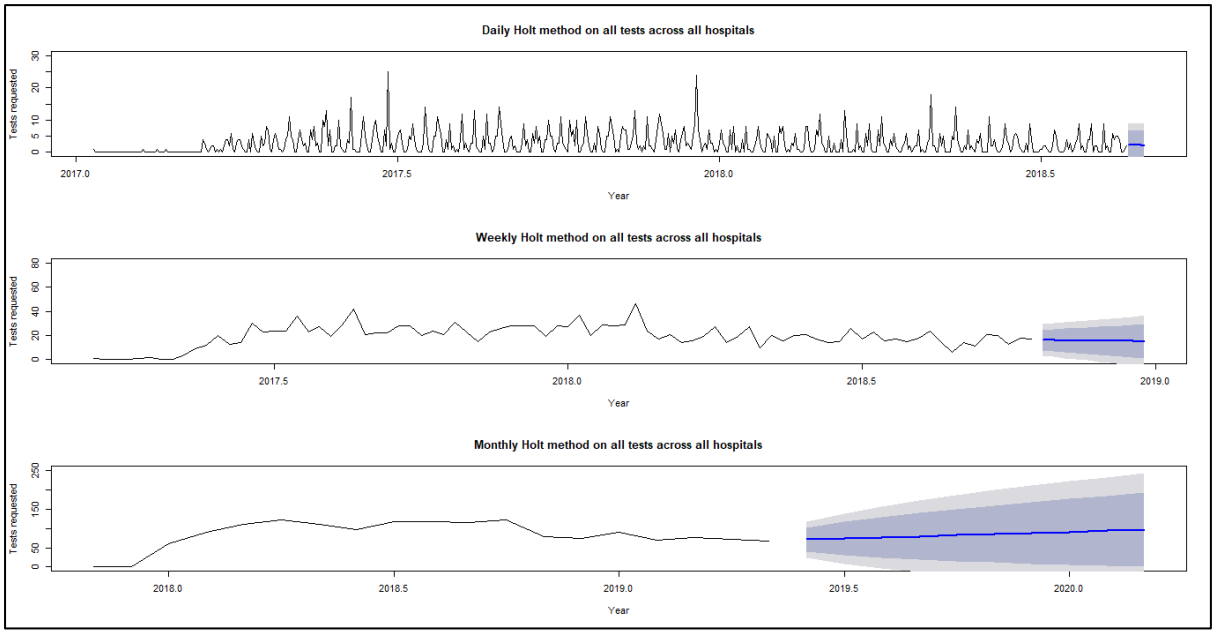

Figure 1.33: Holt-linear- daily/weekly/monthly tests (all hospitals).

These support our suspicion of nonstationarity. From our *ADF tests*, the weekly/monthly *referrals* appear long-term nonstationary, displaying upward trend. We see downward trend in the weekly *tests*, also determined to be nonstationary. There is a slight upward trend in the monthly *tests*, despite a lack of evidence for nonstationary.

The advantage of Holt's method is its accounting for trend. T_t estimates the series trend, therefore Holt's method is more suitable for nonstationary data than naïve and SES methods. However, a limitation of the Holt-linear method is its inability to consider seasonality, instead imposing a linear structure. Our final extrapolation model, *Holt-Winters*, improves Holt-linear by calculating seasonal components.

Holt-Winters method

We conclude our study of extrapolation models with the *Holt-Winters* method, devised by Holt and Winters in 1960[23]. This method captures the seasonal component of series where observations are larger/smaller based on the day/week/month. The forecasting function for the *additive* Holt-Winters method is:

$$\hat{Y}_{t+n} = E_t + nT_t + S_{t+n-p}$$

E_t and T_t are calculated using the previous formulae. We consider n future periods and p seasons in our data. S_t (seasonal component) is given by:

$$S_t = \gamma(Y_t - E_t) + (1 - \gamma)S_{\mathrm{t-p}}$$

γ represents the smoothing parameter for the S_t. Again, *R* optimises γ from minimising SSE. As we examined from decompositions, our series display seasonal components. When fitting the Holt-Winters method on our series, we encounter several issues:

1. Frequency of daily/weekly data is too high: *R* recommends a maximum frequency of 24, as using 52 or 365 seasonal parameters significantly overcomplicates the model. To overcome this, we use XLSTAT to fit Holt-Winters without a frequency.
2. We require 2 periods of data for meaningful results: this posed a significant problem for the monthly datasets as the full dataset contained less than 2 years of observations. Several solutions to this are proposed, including backdating the series as recommended by Chatfield and Yar (1988)[24], using the seasonal naïve method. Instead we use XLSTAT which allows short series but consider that, due to the lack of seasonal data, seasonal components may lack validity.

We produce our forecasts:

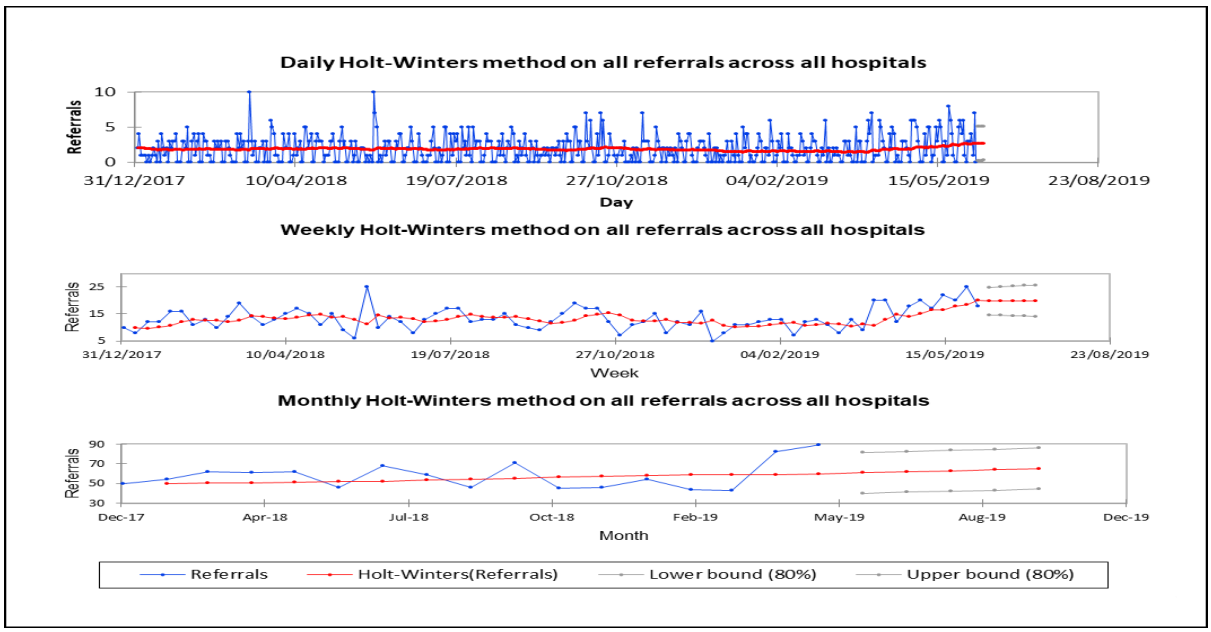

Figure 1.34: Holt-Winters- daily/weekly/monthly referrals (all hospitals) in XLSTAT.

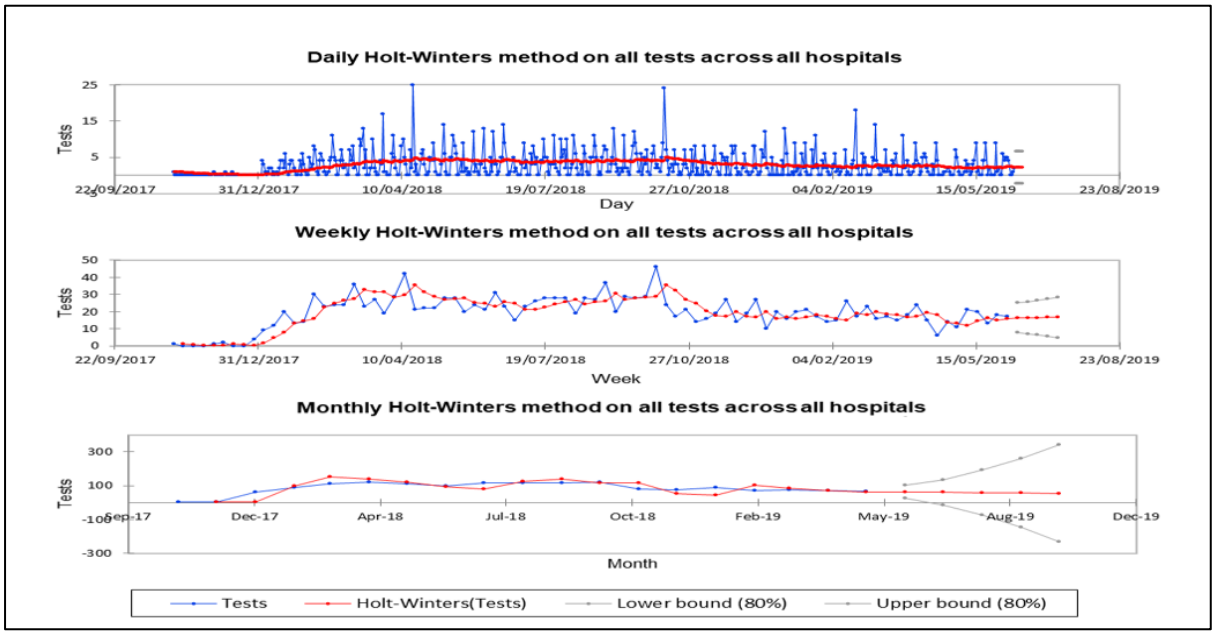

Figure 1.35: Holt-Winters- daily/weekly/monthly tests (all hospitals) in XLSTAT.

The Holt-Winters method perform better for *tests* than *referrals*, presumably as we have longer *tests* data, allowing effective seasonal component calculation.

Holt-Winters is a robust extrapolation model which models level, trend, and seasonality. Many organisations use Holt-Winters for accurate forecasting, its longevity a testament to its effectiveness. However, the method is unsuitable if the seasonal component of our data is

unclear or data is scarce. Furthermore, Holt-Winters is susceptible to outliers[25] and the method does not account for sudden changes due to external factors.

1.4.4 Causal models

We model the data using *causal models* (*simple*/*multiple linear regression*). These methods model the relationship between a response variable and explanatory variable(s). Such models assume linearity between variables and use the least-squares method for fitting.

Simple linear regression (SLR)

SLR functions have form:

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t$$

y_t equals some constant β_0 plus some coefficient β_1 multiplied by a corresponding x_t value plus a random error component ϵ_t[26]. We assume error terms ϵ_t are independent and identically distributed (i.i.d) with normal distribution and mean zero. We apply the model to our *referrals/tests* datasets.

We calculate β_0 (intercept), and coefficient β_1:

Table 1.9: SLR- coefficients for referrals (all hospitals).

Referrals: SLR	Intercept (β_0)	Coefficient (β_1)
Daily		
All referrals, all hospitals	-11.02	0
Weekly		
All referrals, all hospitals	-80.85	0.01
Monthly		
All referrals, all hospitals	-286.61	0.02

We fit our model:

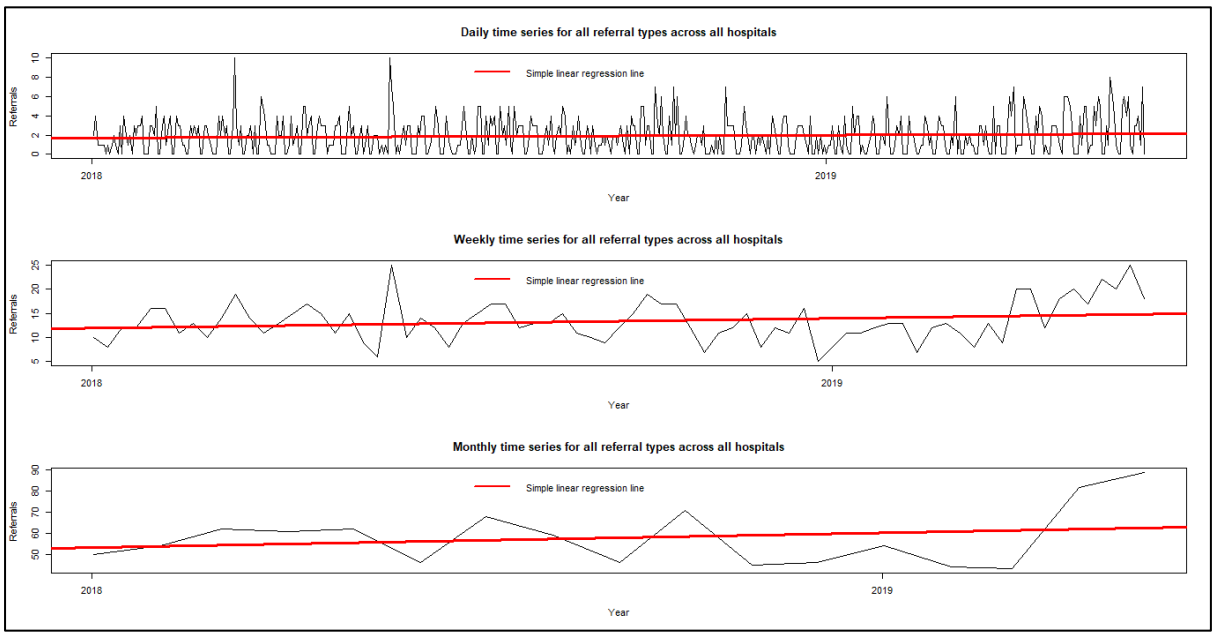

Figure 1.36: SLR- daily/weekly/monthly referrals (all hospitals).

We consider *tests*:

Table 1.10: SLR- coefficients for tests (all tests).

Tests: SLR	Intercept (β_0)	Coefficient (β_1)
Daily		
All referrals, all hospitals	-21.43	0
Weekly		
All referrals, all hospitals	-141.50	0.01
Monthly		
All referrals, all hospitals	-719.63	0.05

We fit our model:

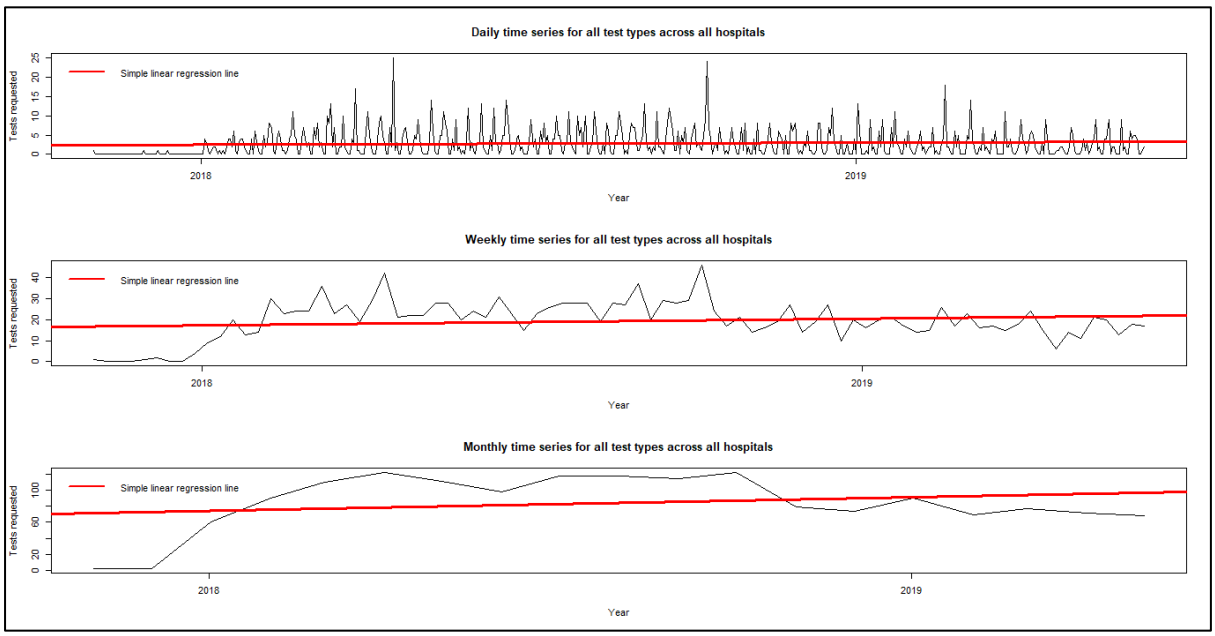

Figure 1.37: SLR- daily/weekly/monthly tests (all tests).

The large negative coefficients from these models results from how *R* treats dates as numeric values (dates are stored as days since 01/01/1970). We discuss the benefits and limitations of SLR.

The main advantage of SLR is its simplicity, it is easy to implement and interpret and useful when we suspect linearity between variables. However, the method is prone to underfitting. Jabbar & Khan (2014) argue that SLR tends to underfit complex data with no clear linear structure[27]. For such series, we desire a more complex model to capture variability. As our *referrals/tests* series have a broadly nonlinear structure, this method may not capture all variability.

Multiple linear regression (MLR)

MLR examines the impact of several explanatory variables on some response variable, taking form:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t} + \epsilon_t$$

[28]for k predictor variables. We add three new variables to the dataset: weekday, month, year. This model determines which explanatory variables have a statistically significant effect on the response variable and how much they affect its value. We focus on daily data:

Table 1.11: MLR- coefficients for referrals (all hospitals).

Referrals: multiple linear regression	Estimate Std.	Error	*t-value*	*p-value*	Statistical significance (Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)
Intercept	1.84	0.26	7.17	1.91e-12	***
Day					
Monday	*Default*				
Tuesday	0.32	0.22	1.46	0.15	
Wednesday	0.51	0.22	2.29	0.02	*
Thursday	0.24	0.22	1.06	0.29	
Friday	0.11	0.22	0.49	0.63	
Saturday	-2.31	0.22	-10.39	< 2e-16	***
Sunday	-2.41	0.22	-10.80	< 2e-16	***
Month					
January	*Default*				
February	0.26	0.29	0.88	0.38	
March	0.19	0.28	0.66	0.51	
April	0.77	0.29	2.70	0.01	**
May	0.91	0.28	3.20	0	**
June	0.60	0.29	2.11	0.04	*
July	1.10	0.28	3.89	0	***
August	0.47	0.28	1.66	0.10	.
September	0.22	0.29	0.78	0.44	
October	0.68	0.28	2.40	0.02	*
November	0.52	0.29	1.81	0.07	.
December	0.28	0.35	0.80	0.42	
Year					
2018	*Default*				
2019	0.42	0.12	3.44	0	***

R takes the first value alphabetically for each variable and compares coefficients for the variable against it. Putting an asterisk (*) before *Monday, January* and *TwentyEighteen* means these become base values for *day, month, year*, respectively. All coefficients are therefore calculated from these base values.

Strong statistical evidence suggests lower weekend referrals, with many on Wednesdays. Similarly, there are more referrals in spring/summer. This confirms our conjecture from boxplot analysis of heightened spring/summer activity and reduced weekend activity. We see statistically significant increases in referrals from 2018-2019, shown by our plots. Next we perform this analysis on our *tests*, setting our earliest year (2017) as base value:

Table 1.12: MLR- coefficients for tests (all tests).

Tests: multiple linear regression	Estimate Std.	Error	*t-value*	*p-value*	Statistical significance (Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)
Intercept	0.64	0.60	1.07	0.29	
Day					
Monday	*Default*				
Tuesday	-0.33	0.37	-0.90	0.37	
Wednesday	4.58	0.37	12.36	< 2e-16	***
Thursday	-0.58	0.37	-1.58	0.11	
Friday	-1.08	0.37	-2.91	0	**
Saturday	-3.04	0.37	-8.24	1.24e-15	***
Sunday	-3.08	0.37	-8.34	5.82e-16	***
Month					
January	*Default*				
February	0.68	0.44	1.54	0.12	
March	0.98	0.43	2.28	0.02	*
April	0.88	0.43	2.03	0.04	*
May	0.67	0.43	1.55	0.12	
June	0.94	0.49	1.93	0.05	.
July	1.04	0.54	1.92	0.06	.
August	1.26	0.54	2.32	0.02	*
September	1.30	0.55	2.38	0.02	*
October	1.31	0.54	2.42	0.02	*
November	0	0.50	-0.01	0.99	
December	0.06	0.50	0.11	0.91	
Year					
2017	*Default*				
2018	2.43	0.44	5.59	3.56e-08	***
2019	1.69	0.51	3.32	0	***

Interestingly, of all days with 10+ test requests, 87% were Wednesdays. We encourage further investigation into why test requests are so common on Wednesdays and uncommon on Fridays/weekends. More test requests are processed during spring/summer. This seasonal variation should also prompt further investigation. There is yearly variation; test requests increase from 2018-2019.

MLR examines the causal relationship between explanatory variables and the response variable. It is also used for forecasting, using the coefficients from each explanatory variable to forecast future values. It is an effective forecasting tool for multivariate data, however, is computationally expensive[21].

MLR: deprivation and lung cancer

As an interesting aside, we examine the relationship between deprivation and lung cancer incidence. Deprivation is a key factor in lung cancer rates according to various studies in Scotland[29] and England[2]. We examine the dataset for cancer rates in Wales[30]. Performing MLR, we examine the relationship between lung cancer and deprivation by analysing crude cancer rates (all cancers):

Table 1.13: MLR- sex and deprivation on Welsh cancer rates.

Crude cancer rates: multiple linear regression	Estimate Std.	Error	*t-value*	*p-value*	Statistical significance (Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)
Intercept	74.32	6.77	10.99	< 2e-16	***
Sex					
Men	*Default*				
Women	-25.93	5.48	-4.73	2.37e-06	***
Deprivation					
Least deprived	*Default*				
Next least deprived	2.16	8.66	0.25	0.80	
Middle deprived	3.94	8.66	0.46	0.65	
Next most deprived	6.98	8.66	0.81	0.42	
Most deprived	11.47	8.66	1.33	0.19	

Being male appears a statistically significant risk factor in developing cancer. While cancer rates increase with deprivation, it is not statistically significant. We compare to lung cancer only:

Table 1.14: MLR- sex and deprivation on Welsh lung cancer rates.

Crude lung cancer rates: multiple linear regression	Estimate Std.	Error	*t-value*	*p-value*	Statistical significance (Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)
Intercept	73.30	1.58	46.32	< 2e-16	***
Sex					
Men	*Default*				
Women	-40.07	1.29	-31.02	< 2e-16	***
Deprivation					
Least deprived	*Default*				
Next least deprived	14.55	2.04	7.12	7.53e-11	***
Middle deprived	27.34	2.04	13.38	< 2e-16	***
Next most deprived	48.77	2.04	23.87	< 2e-16	***
Most deprived	75.99	2.04	37.20	< 2e-16	***

Being male is also a risk factor for developing lung cancer. We also see drastic increases in the likelihood of developing lung cancer based on deprivation. This suggests that deprivation levels are an important factor for lung cancer incidence. This supports the assertion that lung cancer is the cancer type with the strongest association with deprivation of all common cancers[31]. The incidence of lung cancer is twice as high in the most deprived areas as in the least deprived areas of Wales[31]. Higher incidence in deprived areas may be attributable to environmental tobacco smoke and silicosis from coal mining.

The region served by CTMUHB has among the highest rates of lung cancer in Wales[31]. Smoking in this region is more common than the Welsh average[32], and the region has a large historical coal mining industry. Further research should investigate the relationship between deprivation, smoking and coal mining in South Wales. We return to time series analysis by discussing ARIMAs.

1.4.5 ARIMAs

The ARIMA(*p,d,q)* model consists of three terms:

p – order of autoregressive (AR) term.
q – order of moving average (MA) term.
d – differencing required to achieve stationarity[33].

We consider the simpler *autoregressive moving average* (ARMA) model. The autoregressive component of the model regresses on previous observation values, whereas the moving average component models the error value using a linear combination. We define the ARMA(p,q) model:

$$\hat{Y}_t = c + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-1} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t$$

[34]for characteristic polynomial for AR component $\Phi(y)$ and characteristic polynomial for MA component $\theta(y)$. This formula becomes an ARIMA by differencing.

We select an ARIMA model by examining the *autocorrelation function* (ACF) (the relationship between current and historic values). We also consider *partial autocorrelation function* (PACF) (the relationship between current and historic values, omitting intervening observations). These help us determine which ARIMA to select:

Table 1.15: characteristics of ARMA models.

Model	ACF	PACF
AR(p)	Dies away geometrically	Zero after lag p
MA(q)	Zero after lag q	Dies away geometrically
ARMA(p, q)	Dies away geometrically	Dies away geometrically

We consider if our series requires differencing (transforming non-stationary series into stationary ones), using our Dickey-Fuller test. Differencing is required for suspected non-stationary series, including weekly/monthly *referrals* (all hospitals) and weekly *tests* (all tests).

To find an appropriate AR term, we consider our PACFs:

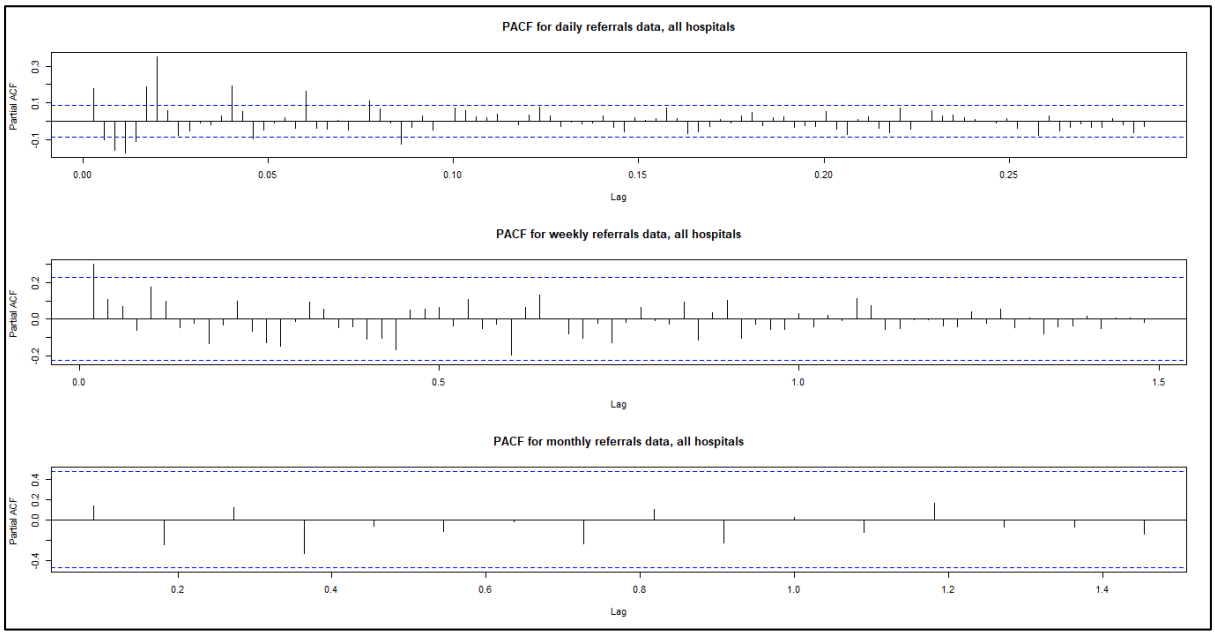

Figure 1.38: PACF- daily/weekly/monthly referrals (all hospitals).

We decide our AR term by counting lags which fall outside our significance region (we may allow several outliers). We select AR(7), AR(1), AR(0) for daily, weekly, monthly *referrals* respectively as all values fall within the significance region.

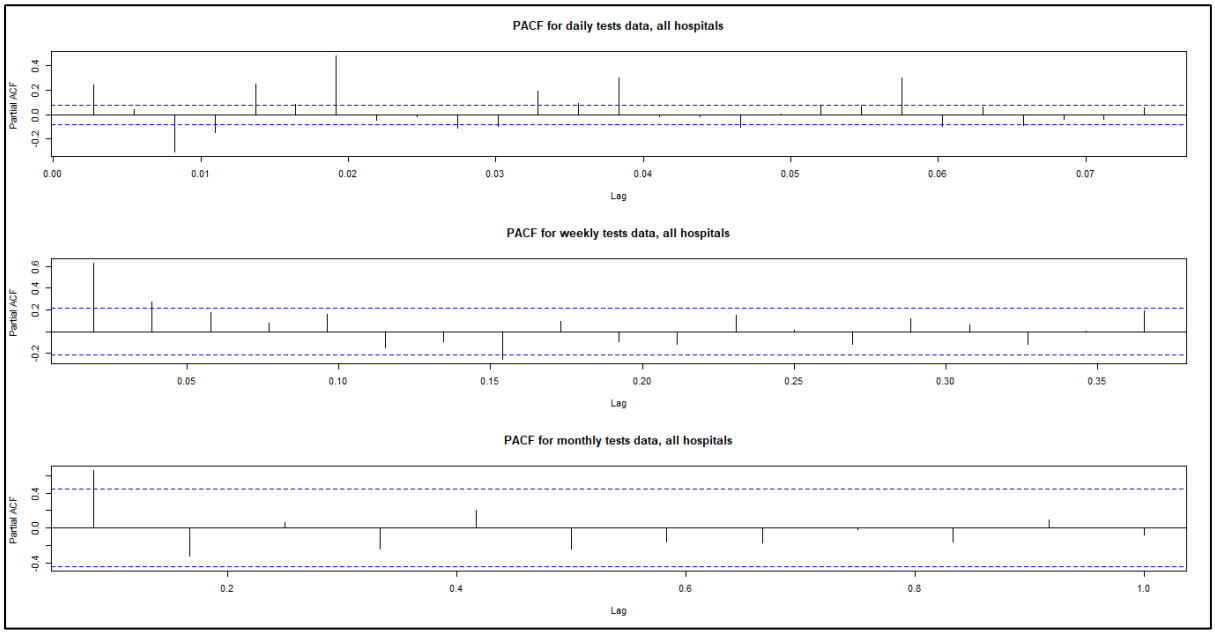

Figure 1.39: PACF- daily/weekly/monthly tests (all tests).

We next consider the MA(q) term, estimated from the ACF for our series.

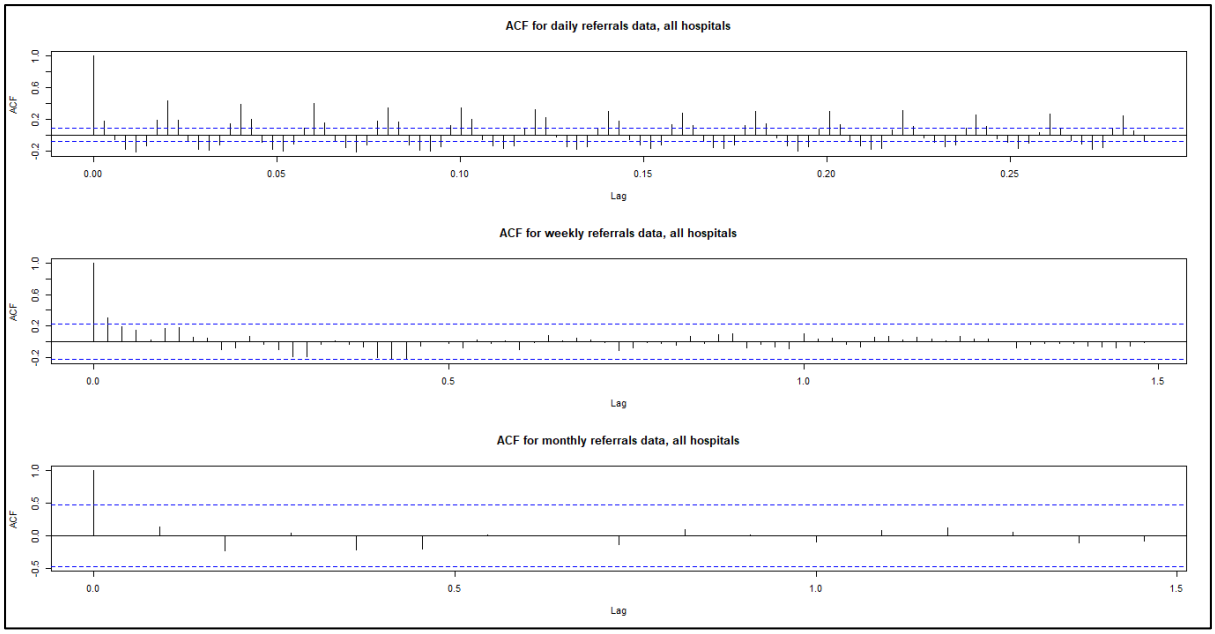

Figure 1.40: ACF- daily/weekly/monthly referrals (all hospitals).

The ACF for our daily data dies away geometrically, therefore we suspect that an ARIMA(7,0,7) model is suitable. The ACF for our weekly data dies away after lag 1, therefore we choose ARIMA(1,1,1). Finally, the ACF for our monthly data is zero after lag 0, therefore we select ARIMA(0,1,0). Finally, we assess the ACF for our *tests* series:

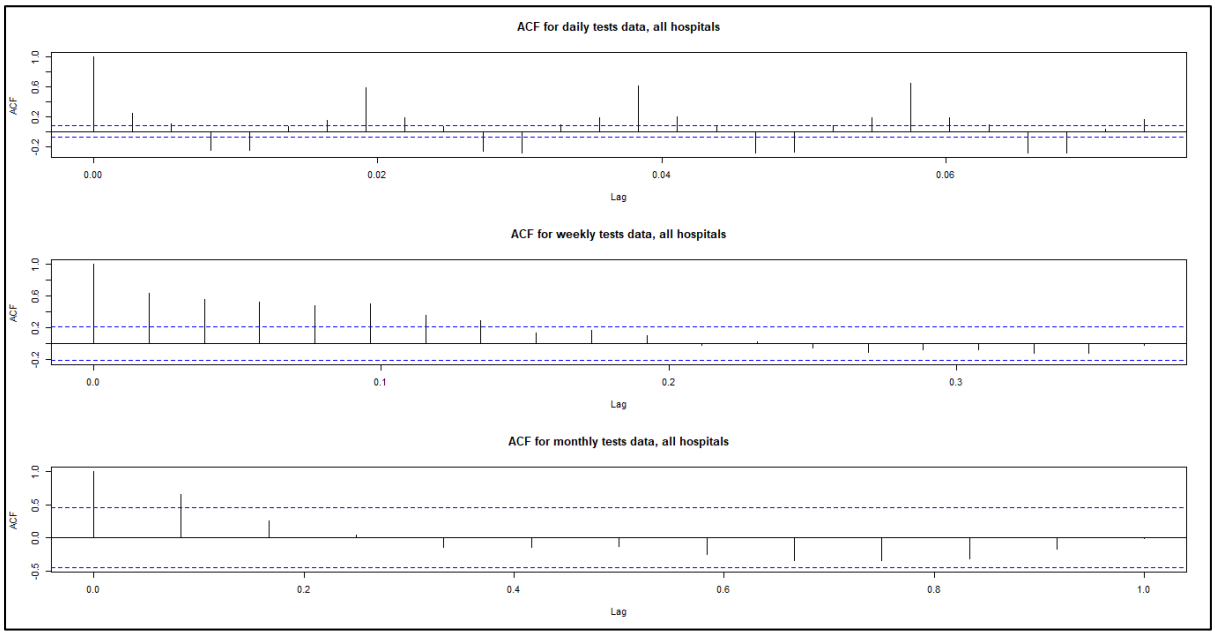

Figure 1.41: ACF- daily/weekly/monthly tests (all tests).

The ACF for our daily data dies away geometrically, suggesting an ARIMA(7,0,7). The weekly data also dies away geometrically, suggesting an ARIMA(1,1,1). Lastly, the monthly series' ACF drops after lag 1, suggesting an ARIMA(0,1,1).

The *tsdiag* function runs diagnostic tests on ARIMA models. The function plots standardised residuals, residuals' ACFs and *p*-values of *Ljung-Box statistics* (tests whether autocorrelations are nonzero). For a well-fitting model, we require residuals to be independent and identically distributed with mean zero. We also require the ACF of residuals to fall within some confidence interval and the Ljung-Box statistics to be above p-value 0.05.

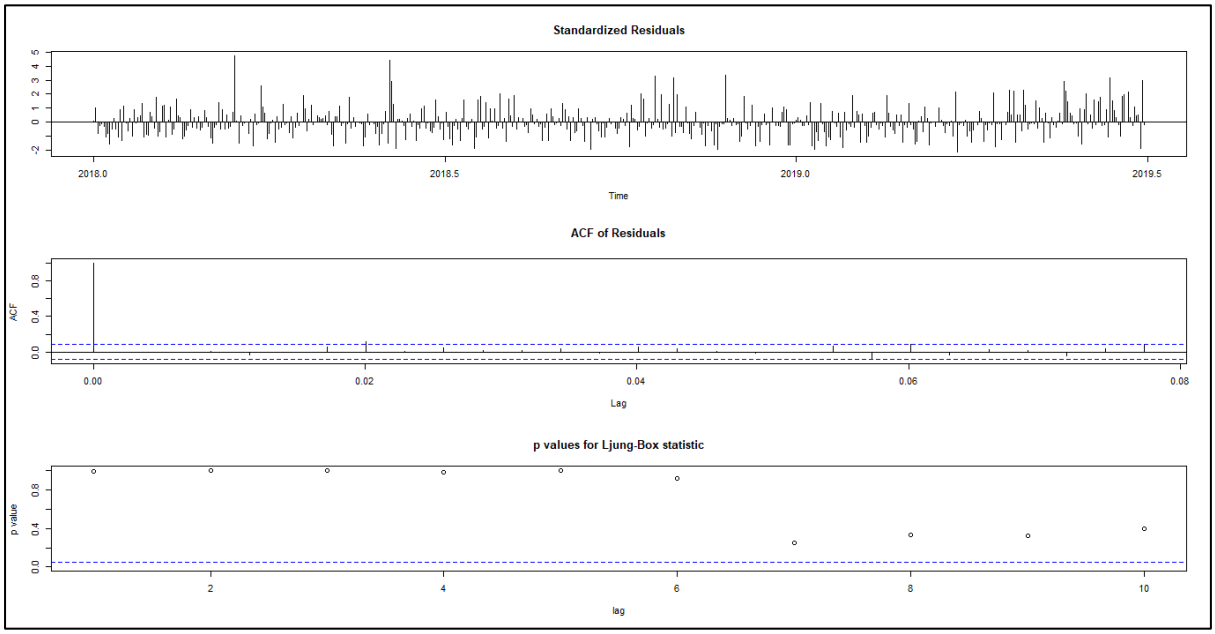

Figure 1.42: diagnostic test- ARIMA(7,0,4) on daily referrals (all hospitals).

Such diagnostic tests on suggested ARIMA models reveal that all models are suitable. We finally consider the *principle of parsimony* (endorsing the simplest suitable model which passes our diagnostics tests)[35]. We endorse:

Table 1.16: most parsimonious ARIMA models.

Time series	Parsimonious ARIMA model
Referrals (all hospitals)	
Daily	(7,0,4)
Weekly	(1,1,1)
Monthly	(0,1,0)
Tests (all tests)	
Daily	(7,0,7)
Weekly	(1,1,1)
Monthly	(0,0,1)

We apply these models:

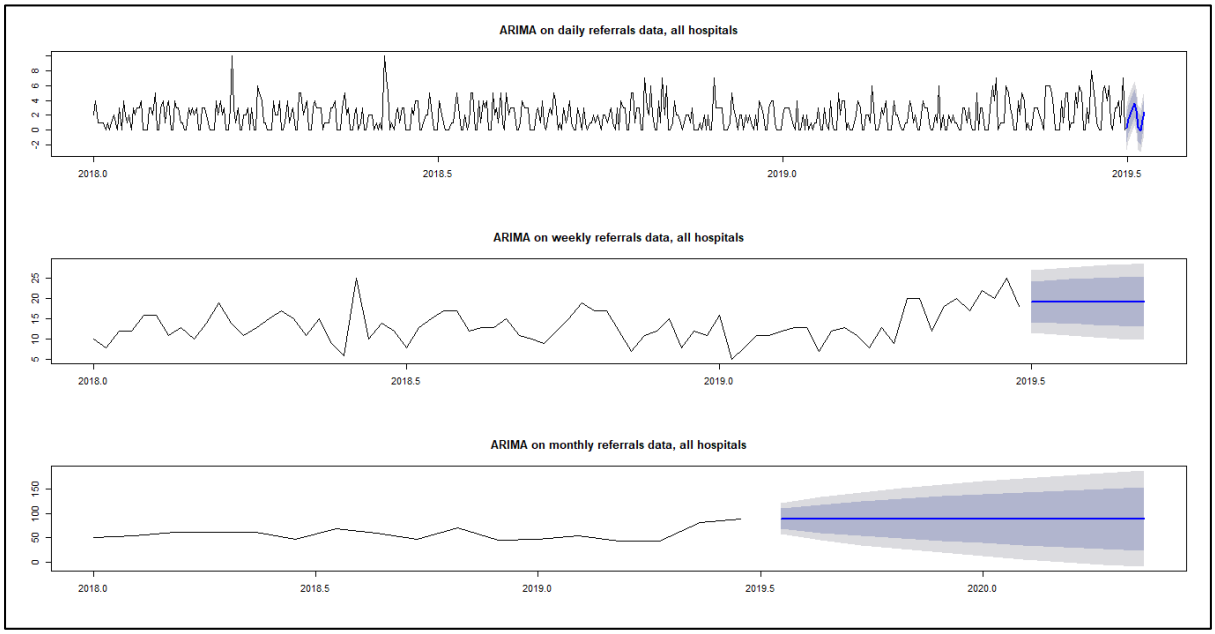

Figure 1.43: ARIMA- daily/weekly/monthly referrals (all hospitals).

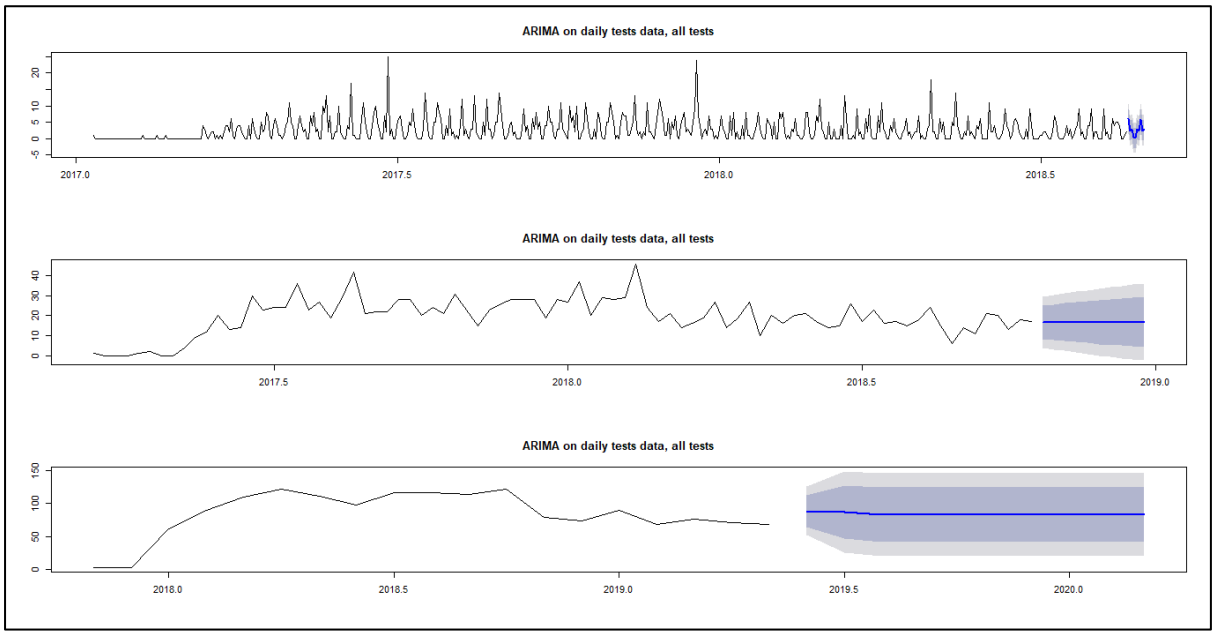

Figure 1.44: ARIMA- daily/weekly/monthly tests (all tests).

ARIMA models rely on the forecaster to select an appropriate model. Furthermore, because ARIMA models predict based on previous values, they often fail to identify turning points.

1.4.6 Singular spectrum analysis (SSA)

SSA is a new time series analysis method, combining fields including classical time series analysis, multivariate statistics and multivariate geometry[36]. SSA expresses *trend*, *seasonality*, and *residuals* as sets of components. This method can reconstruct the time series using some subset of components, excluding unwanted noise. We isolate components, examining their structure, including noise components. For a time series of length N, $x_1, \dots, x_N$, we select window length L ($1 < L < N$) and construct vectors $X_i = (x_i, \dots, x_{i+L-1})$ where $i = 1, \dots, K$. We combine these vectors into a matrix $\boldsymbol{X} = (x_{i+j-1})_{i,j=1}^{L,K}$. We perform *singular value decomposition* (SVD) on $\boldsymbol{X}\boldsymbol{X}^T$ to extract L *eigenvectors* and *eigenvalues*. Taking some combination of $l < L$ eigenvectors, the set of vectors $\{X_1, \dots, X_K\}$ is projected onto an l-dimensional subspace. After averaging over the diagonals of our matrix, we obtain $\bar{\boldsymbol{X}}$, which may be used as an approximation for $\boldsymbol{X}$[37]. Our approximation $\bar{\boldsymbol{X}}$ allows time series forecasting.

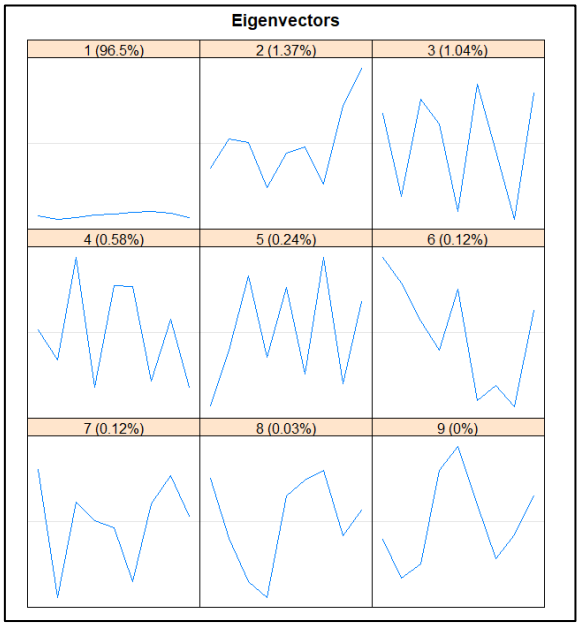

Figure 1.45: eigenvectors (and the amount of variation they account for)- monthly referrals (all hospitals) training data.

We determine the eigenvectors which compose our time series and assess which account for the most variability within our data.

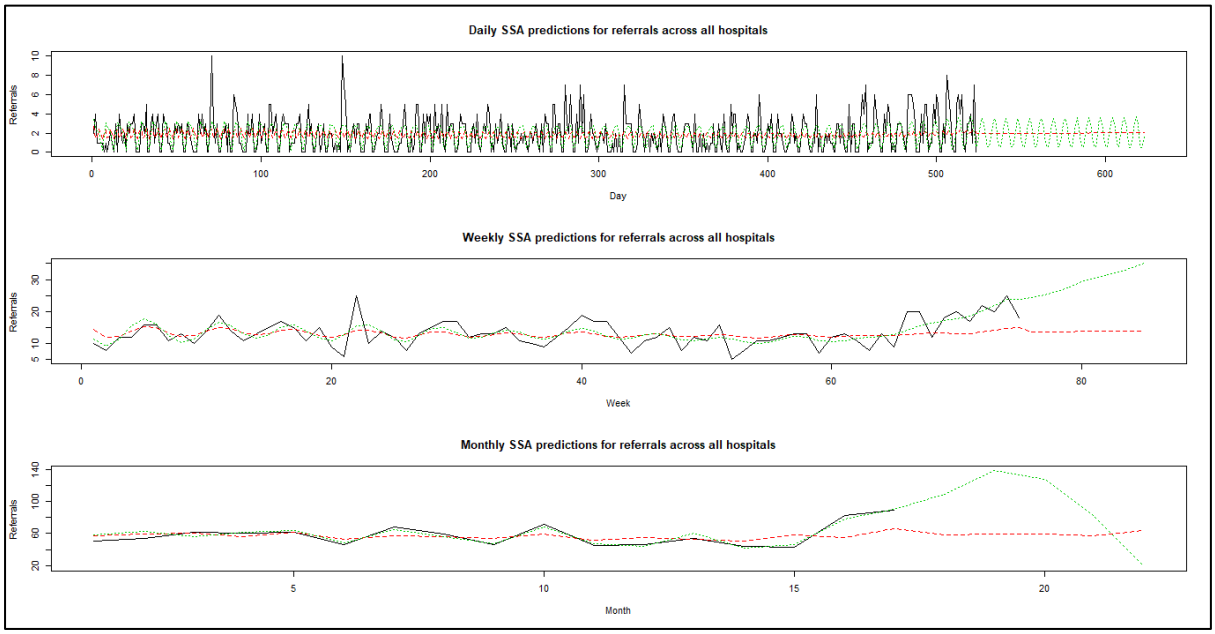

Figure 1.46: SSA decomposition- referrals (all hospitals) data trend (red line) and noise (green line).

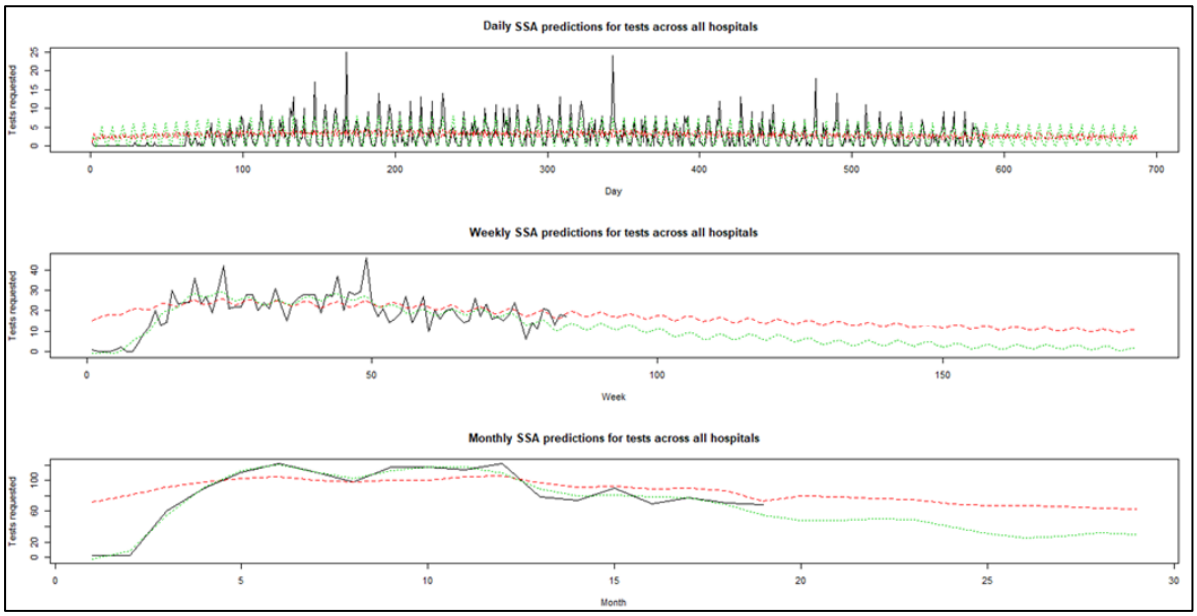

Figure 1.47: SSA decomposition- tests (all tests) data trend (red line) and noise (green line).

We forecast using SSA:

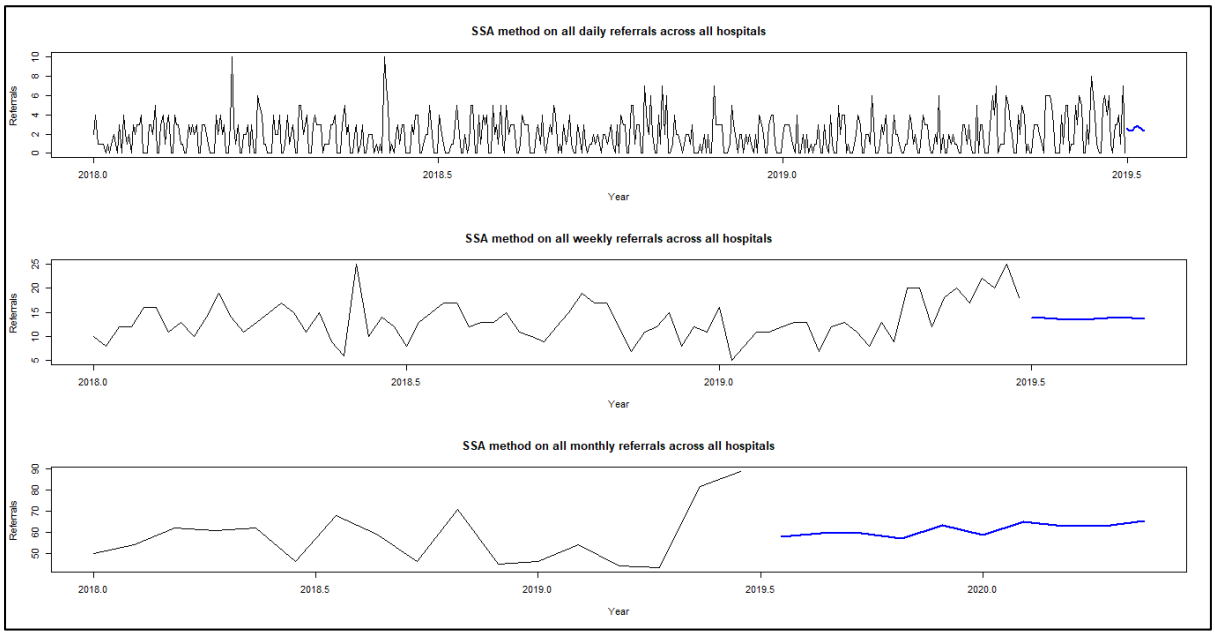

Figure 1.48: SSA- daily/weekly/monthly referrals (all hospitals).

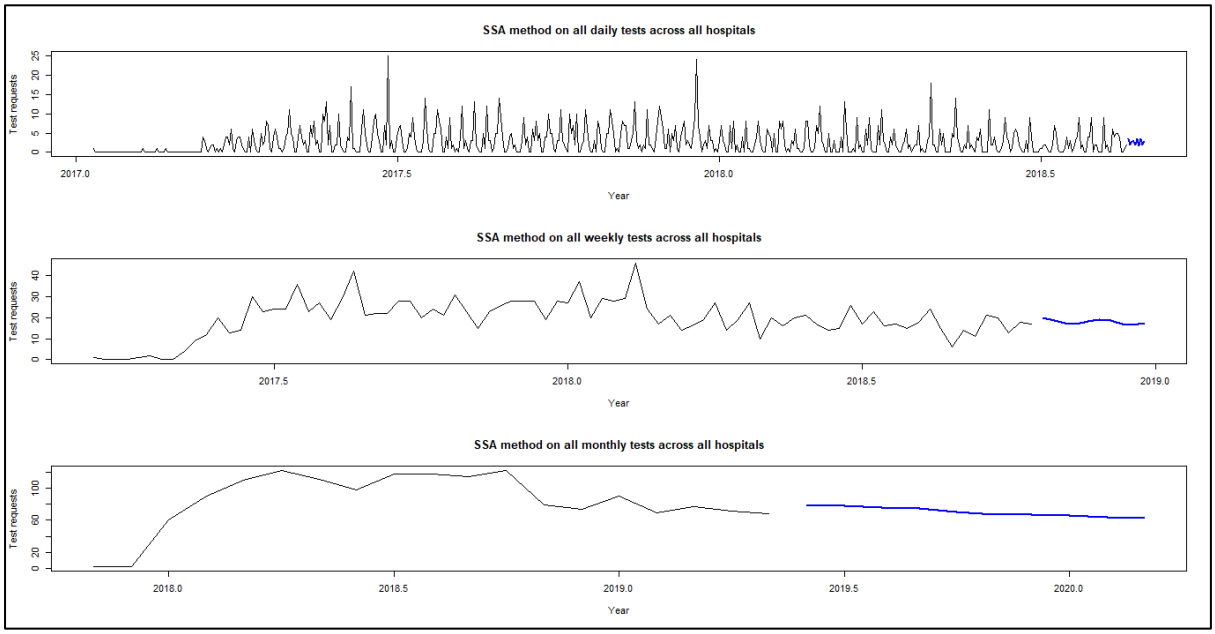

Figure 1.49: SSA- daily/weekly/monthly tests (all tests).

SSA allows us to examine each component of the decomposition of a series and potentially remove noise. It has been proven as an effective forecasting method in many scientific areas[38]. Unlike ARIMAs, SSA does not use assumptions from the series' structure, instead relying on a single parameter[39]. The main disadvantage of SSA is its computational complexity. Calculating the SVD of matrices requires large amounts of memory, therefore

this method is unsuitable for complex series. We next discuss a machine learning technique for forecasting: ANNs.

1.4.7 Machine learning- ANNs

ANNs are computational models, inspired by the human brain. In human biology, neurons receive signals through synapses. Such neurons activate should the signal exceed some threshold. We model synapses as inputs, with signal strength given some weight value. The neuron is potentially activated by some function. We then calculate the output for these neurons and send it to the rest of the network.

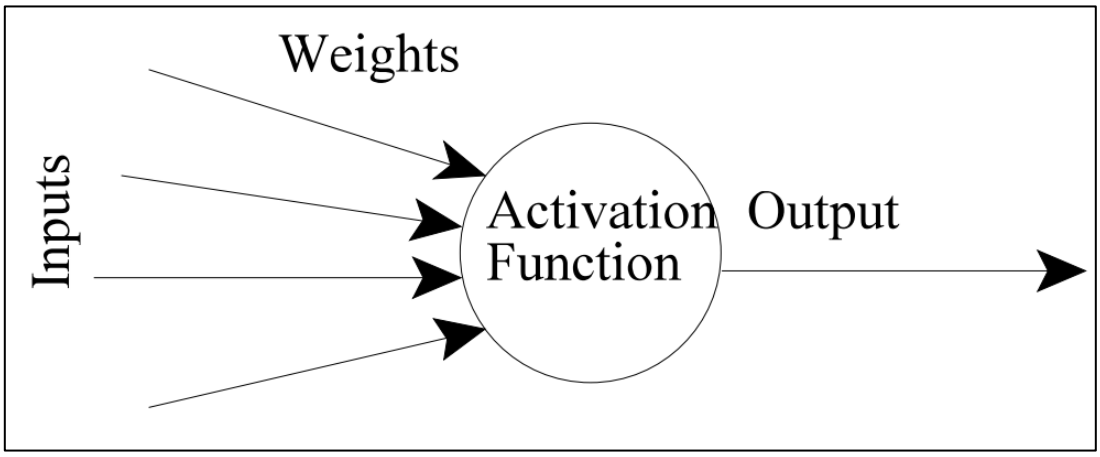

Figure 1.50: artificial neuron- inputs, weights, activation function and output[40].

We multiply each input x by its weight w:

$x_1 \rightarrow x_1 * w_1$
$x_2 \rightarrow x_2 * w_2$
...
$x_n \rightarrow x_n * w_n$

Summing all weighted inputs and inputting some bias parameter γ:

$$a = \gamma + \sum_{i=1}^{n} (x_i * w_i)$$

inputting a into the activation function y:

$$y = f(a)$$

This calculation determines whether a neuron is activated, assigning a value to the output. Combining many artificial neurons creates an ANN, these can be trained to forecast time series.

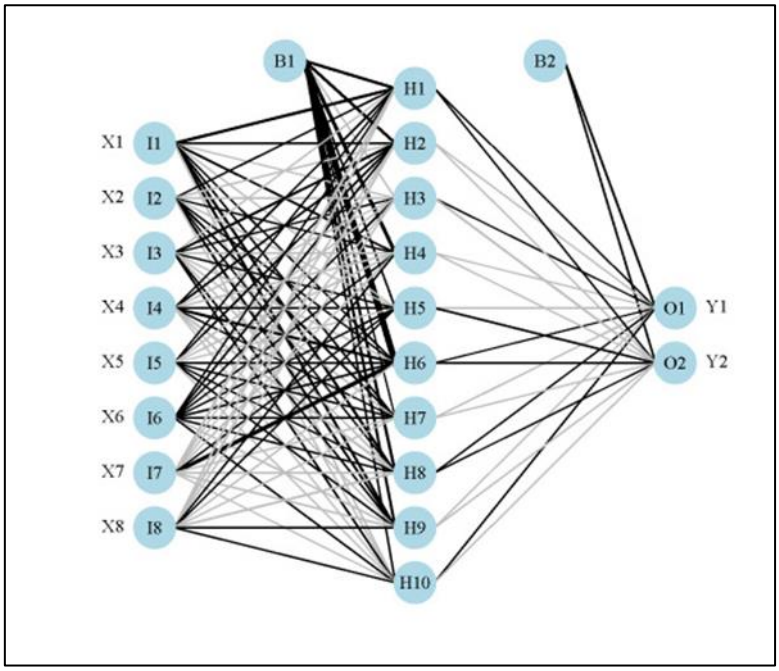

Figure 1.51: an ANN[41].

In the above diagram, '*I'* neurons represent the input section, *'H'* represents the middle section, '*O*' represents the output and '*B'* represents the bias parameters[41]. ANNs are called

'black box' approaches; studying the network's structure does not give insight into the link between weights and functions.

We use the *nnetar* function. This function fits an ANN model to a time series where inputs are lagged time series values.

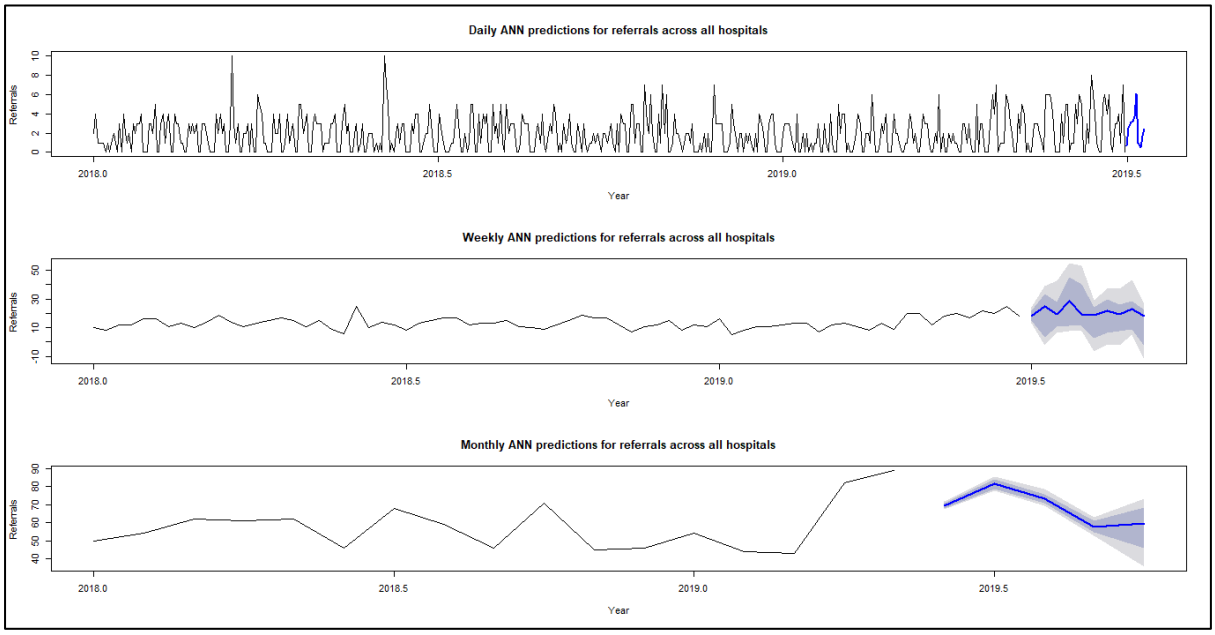

Figure 1.52: ANN- daily/weekly/monthly referrals (all hospitals).

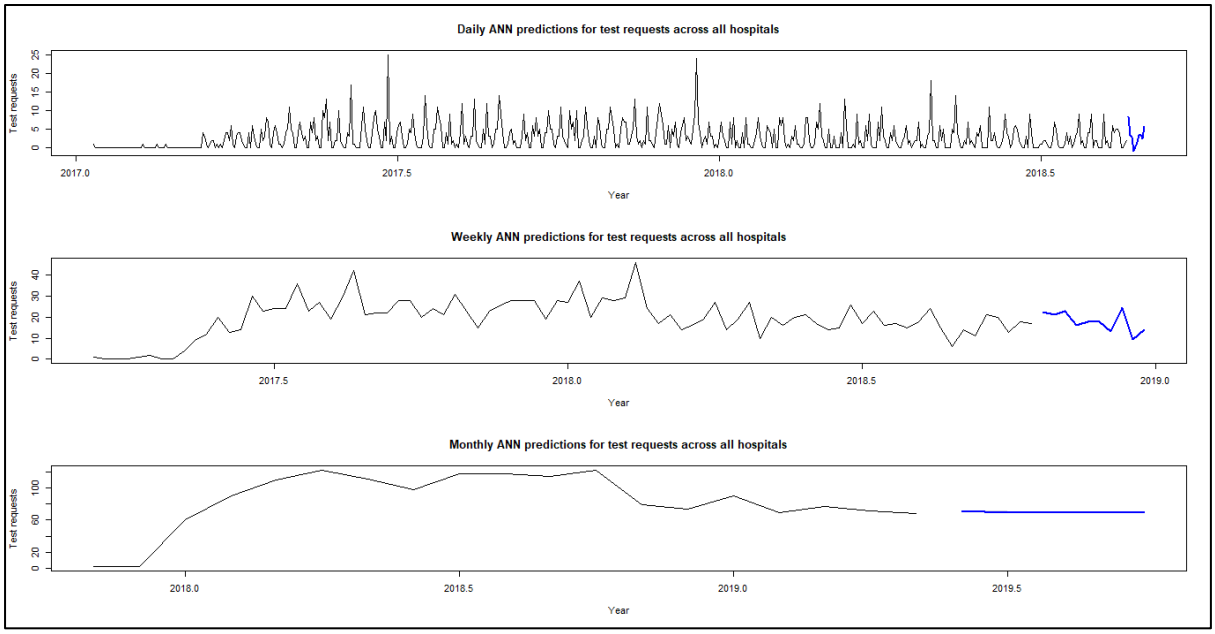

Figure 1.53: ANN- daily/weekly/monthly tests (all tests).

The main advantage of ANNs is their accuracy. Empirical studies vouch for the effectiveness of such models, compared to extrapolation and ARIMA models[42]. ANNs are robust against

noise and can even effectively forecast from incomplete data. Such methods are therefore suitable for real-world imperfect data[43]. ANNs, however, have limitations. As discussed, such methods are 'black boxes'; we struggle to understand causal relationships when using ANNs. Furthermore, due to the complex 'hidden' middle section, ANNs require large computational resources[44].

1.5 Results

To assess the accuracy of each model, we consider error metrics. The magnitude of our error statistics is inversely proportional to the accuracy of the model.

1.5.1 Error statistics

We begin with the mean absolute error (MAE), the simplest error metric used in forecasting. We calculate the sum of the absolute value for all residuals (the difference between some observed value and the forecasted value) for every point in the test set and find the mean error. We define MAE:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

for n datapoints, observed value y_i and forecasted value $\hat{y}_i$[45]. The MAE does not identify whether the model overestimates/underestimates the observed values. We next examine *mean squared error* (MSE), which squares the difference between observed and forecasted values:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

[45]squaring residuals means that extreme values contribute a larger proportion of the error; models are heavily penalised for outliers. The MSE is equivalent to the variance of residual values. The square root of the MSE gives the *root mean squared error* (RMSE): the standard deviation for residual values. We next examine *mean percentage error* (MPE), which allocates a percentage to the MAE:

$$MPE = \frac{100\%}{n}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{y_i}\right)$$

[45]The advantage of the MPE metric is that its output informs us whether our model overfits (positive value) or underfits (negative value) the data. Finally, we examine *mean absolute percentage error* (MAPE):

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

[45]this metric is robust against extreme values as it calculates the absolute value.

Each error metric provides different information about our residuals. Therefore, we decide which metric to use based on the nature of the data. The main limitation of both percentage metrics is their use of the division operator, which causes issues for zero values as division by zero is undefined. As our daily/weekly data often contains zeroes, these error statistics are problematic. We therefore mainly focus on MAE and the MSE for daily/weekly data. Again, the differences between these values informs us about the effect of outliers.

For brevity, we discuss the results from the daily/weekly/monthly *referrals* at all hospitals and daily/weekly/monthly *tests* of all types. Statistics and forecasts for *referrals* time series by hospital and *tests* by test type are contained in the appendix. We calculate and discuss the error statistics for each method on the training and test set, examining whether the accuracy improves or deteriorates when we consider the test set. This allows us to state our confidence with the selected model. Finally, we select a model from the test set results, by examining forecast plots.

1.5.2 Forecasts/results

We examine forecasts for our daily referrals, using the accuracy function. The function provides error statistics of the fit against the training/test set. We forecast using each method for 10 days/10 weeks/3 months then examine the error metrics.

Table 1.17: forecasts from each method- referrals- all hospitals, daily.

Time series method	Forecast (2019)									
Referrals- all hospitals, daily	08/06	09/06	10/06	11/06	12/06	13/06	14/06	15/06	16/06	17/06
Naïve	0	0	0	0	0	0	0	0	0	0
Seasonal Naïve	0	4	2	1	0	0	0	1	1	3
SES	2.46	2.46	2.46	2.46	2.46	2.46	2.46	2.46	2.46	2.46
Holt-linear	3.46	3.48	3.50	3.52	3.54	3.56	3.59	3.61	3.63	6.65
Holt-Winters	1.79	3.78	4.14	4.39	4.34	4.25	2.03	1.96	3.94	4.31
SLR	2.10	2.10	2.10	2.10	2.11	2.11	2.11	2.11	2.11	2.11
MLR	2.32	2.32	2.31	2.31	2.31	2.31	2.31	2.31	2.31	2.30
ARIMA	0.26	1.45	2.42	2.99	3.60	2.63	0.43	0	1.38	2.38
SSA	2.56	2.38	2.33	2.44	2.65	2.81	2.80	2.64	2.45	2.36
ANN	1.45	1.93	2.60	3.88	3.23	5.23	2.44	0.35	2.04	1.84

Table 1.18: error statistics- daily referrals (all hospitals) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- all hospitals, daily training set					
Naïve	1.76	5.70	2.39	Inf	Inf
Seasonal Naïve	1.66	5.72	2.39	-Inf	Inf
SES	1.53	3.46	1.86	-Inf	Inf
Holt-linear	1.52	3.42	1.85	-Inf	Inf
Holt-Winters	1.09	2.24	1.50	Inf	Inf
SLR	1.53	3.45	1.86	-Inf	Inf
MLR	1.14	2.40	1.55	Inf	Inf
ARIMA	1.10	2.23	1.49	Inf	Inf
SSA	1.49	3.27	1.81	-Inf	Inf
ANN	0.01	0	0.01	Inf	Inf

Table 1.19: error statistics- daily referrals (all hospitals) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- all hospitals, daily test set					
Naïve	2.20	9.60	3.10	-Inf	Inf
Seasonal Naïve	2.20	8	2.83	-Inf	Inf
SES	2	4.83	2.20	10.58	81.29
Holt-linear	2.19	6.51	2.55	38.44	61.75
Holt-Winters	2.23	6.76	2.60	33.03	75.40
SLR	2	4.77	2.18	-4.44	94.94
MLR	2	4.78	2.19	4.73	86.62
ARIMA	1.59	5.09	2.26	-99.72	194.16
SSA	1.99	4.73	2.17	14.19	78.26
ANN	1.72	4.02	2	18.22	84.82

Our daily *referrals* contains many zeroes therefore MPE/MAPE are undefined. We restrict our attention, for daily data, to MAE/RMSE. Furthermore, as the daily numbers are small, any errors appear significant, resulting in large MAE/MAPE. The ANN method has low forecast error statistics, as does the ARIMA model. Low error statistics indicate a good fit, but do not guarantee the lowest post-sample errors as methods with low fitting error statistics may have larger errors for post-sample data[46]. We examine a plot of models with low error statistics to determine which model best fits our data:

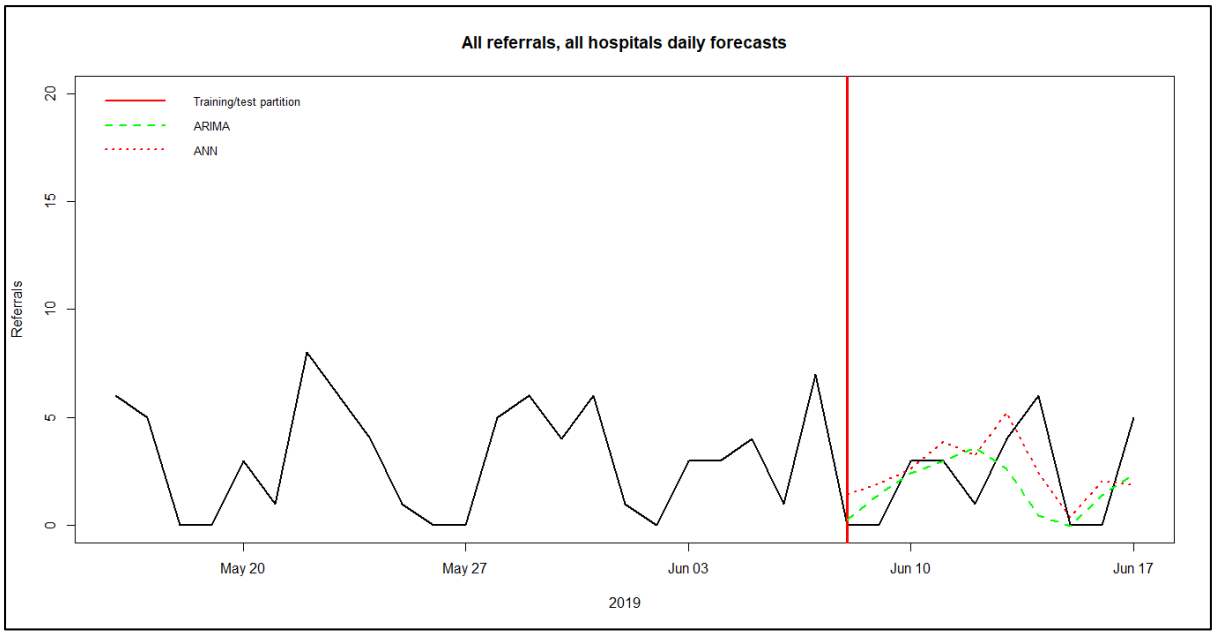

Figure 1.54: forecast- daily referrals (all hospitals).

The ANN method, which scored the lowest error values for the training set, best reflects the structure of the post-sample data. We endorse the ANN model for daily *referrals*. The low

error statistics for the ANN method across the training and test set give confidence in this method. Our ten-day forecast for the entire time series is:

Table 1.20: ten-day forecast- daily referrals (all hospitals).

Date	Forecast (ANN)
01/12/2019	0.74
02/12/2019	2.82
03/12/2019	6.10
04/12/2019	0.96
05/12/2019	2.05
06/12/2019	1.64
07/12/2019	0.83
08/12/2019	0
09/12/2019	3.52
10/12/2019	0.91

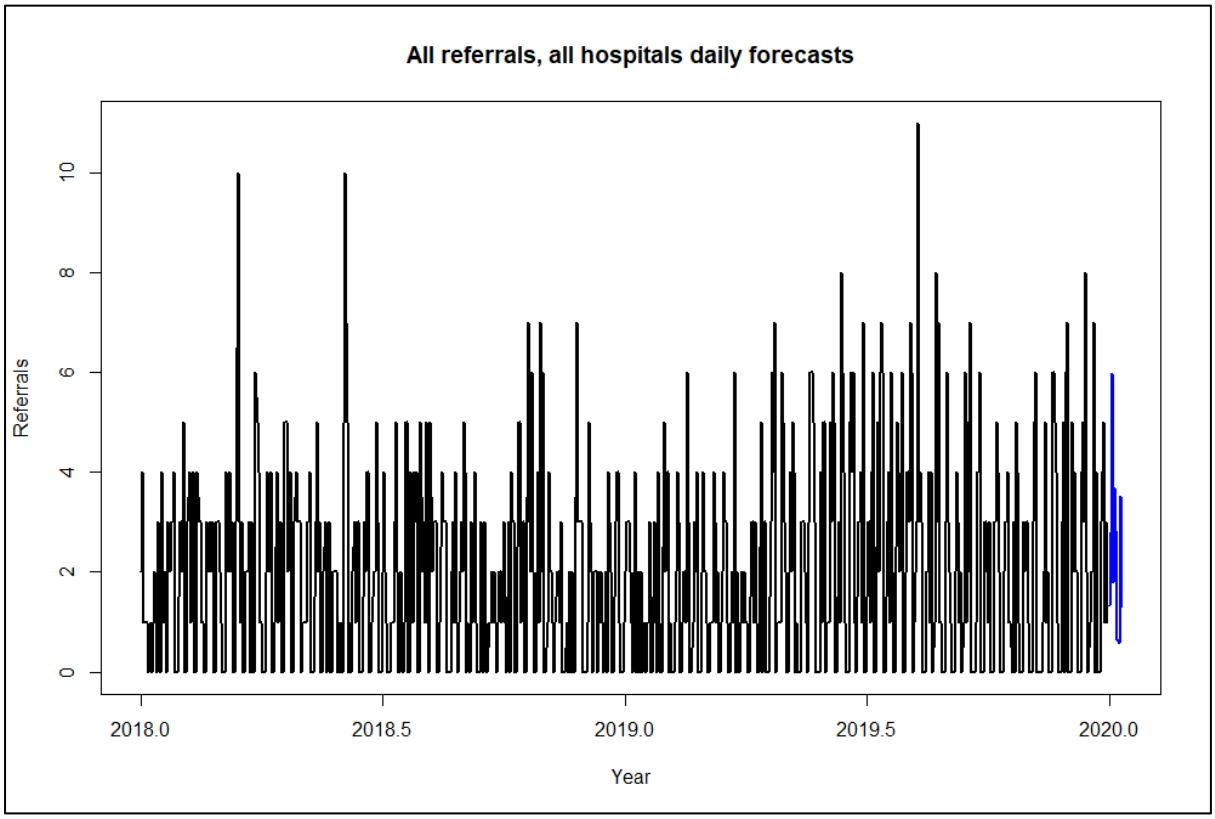

Figure 1.55: forecast- daily referrals (all hospitals).

We examine weekly *referrals* forecasts:

Table 1.21: forecasts from each method- referrals- all hospitals, weekly.

Time series method	Forecast (2019)									
Referrals - all hospitals, weekly	w/c 04/06	w/c 11/06	w/c 18/06	w/c 25/06	w/c 02/07	w/c 09/07	w/c 16/07	w/c 23/07	w/c 30/07	w/c 06/08
Naïve	18	18	18	18	18	18	18	18	18	18
Seasonal Naïve	8	13	15	17	17	17	12	13	15	11
SES	19.38	19.38	19.38	19.38	19.38	19.38	19.38	19.38	19.38	19.38
Holt-linear	24.68	25.74	26.80	27.86	28.92	29.98	31.04	32.10	33.16	34.22
Holt-Winters	18.13	20.91	22.41	19.79	20.81	21.48	22.67	18.04	19.57	19.49
SLR	14.78	14.82	14.86	14.89	14.93	14.97	15	15.04	15.08	15.12
ARIMA	19.23	19.28	19.28	19.28	19.28	19.28	19.28	19.28	19.28	19.28
SSA	13.78	13.83	13.64	13.52	13.57	13.70	13.83	13.87	13.77	13.70
ANN	17.89	24.74	19.61	24.58	20.49	18.09	23.87	18.54	25.78	17.88

And the corresponding error statistics:

Table 1.22: error statistics- weekly referrals (all hospitals) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- all hospitals, weekly					
Naïve	3.65	23.57	4.85	-6.98	31.03
Seasonal Naïve	4.48	30.48	5.52	-3.35	32.54
SES	3.04	15.09	3.88	-5.58	25.55
Holt-linear	3.09	14.40	3.80	-7.30	26.27
Holt-Winters	2.74	12.44	3.53	-6.73	22.23
SLR	3.25	16.31	4.04	-10.39	27.92
ARIMA	3.03	15.08	3.88	-4.94	25.33
SSA	2.99	14.84	3.85	-6.79	24.76
ANN	2.17	8.30	2.88	-5.56	17.33

Table 1.23: error statistics- weekly referrals (all hospitals) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- all hospitals, weekly test set					
Naïve	2.90	12.90	3.59	-7.22	16.11
Seasonal Naïve	6.10	54.50	7.18	-50.38	51.72
SES	2.78	11.22	3.35	0.42	14.32
Holt-linear	10.15	134.22	11.59	33.17	33.17
Holt-Winters	3.02	14.07	3.75	-13	17.83
SLR	4.58	30.53	5.53	-29.20	30.68
ARIMA	2.75	11.21	3.35	-0.13	14.27
SSA	5.58	42.53	6.52	-40.72	40.72
ANN	3.75	23.89	4.89	6.93	17.03

The lowest error statistics for the training set are given by the ANN model, and the lowest error statistics for the test set are given by the ARIMA model. We examine a plot of these models and discuss which to endorse.

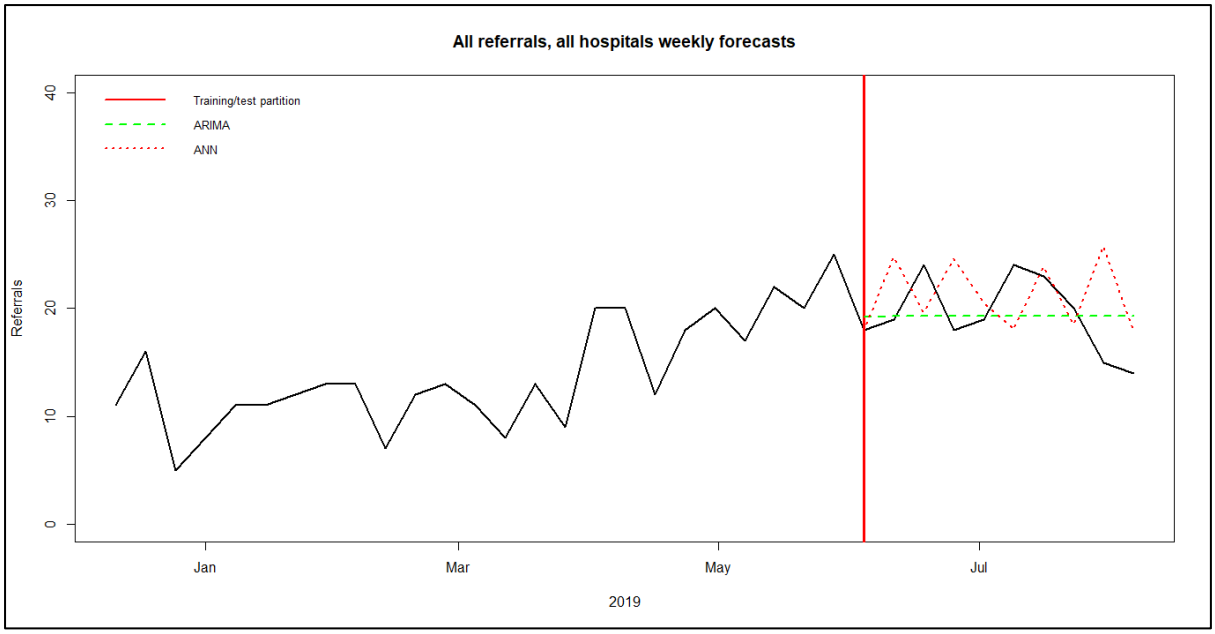

Figure 1.56: forecast- weekly referrals (all hospitals).

Despite scoring low error statistics, the ANN approach overfits the post-sample data. Conversely, the ARIMA forecast underfits the post-sample data and fails to reflect the data's variation. As the ARIMA model has lower error statistics on the test set, we endorse the use of this model for weekly *referrals*, though the relatively large error statistics for this model on the training set reduces our confidence in this model choice. Our ten-week forecast from ARIMA:

Table 1.24: ten-week forecast- weekly referrals (all hospitals).

Week commencing	Forecast (ARIMA)
03/12/2019	16.20
10/12/2019	16.36
17/12/2019	16.37
24/12/2019	16.37
31/12/2019	16.37
07/01/2020	16.37
14/01/2020	16.37
21/01/2020	16.37
28/01/2020	16.37
04/02/2020	16.37

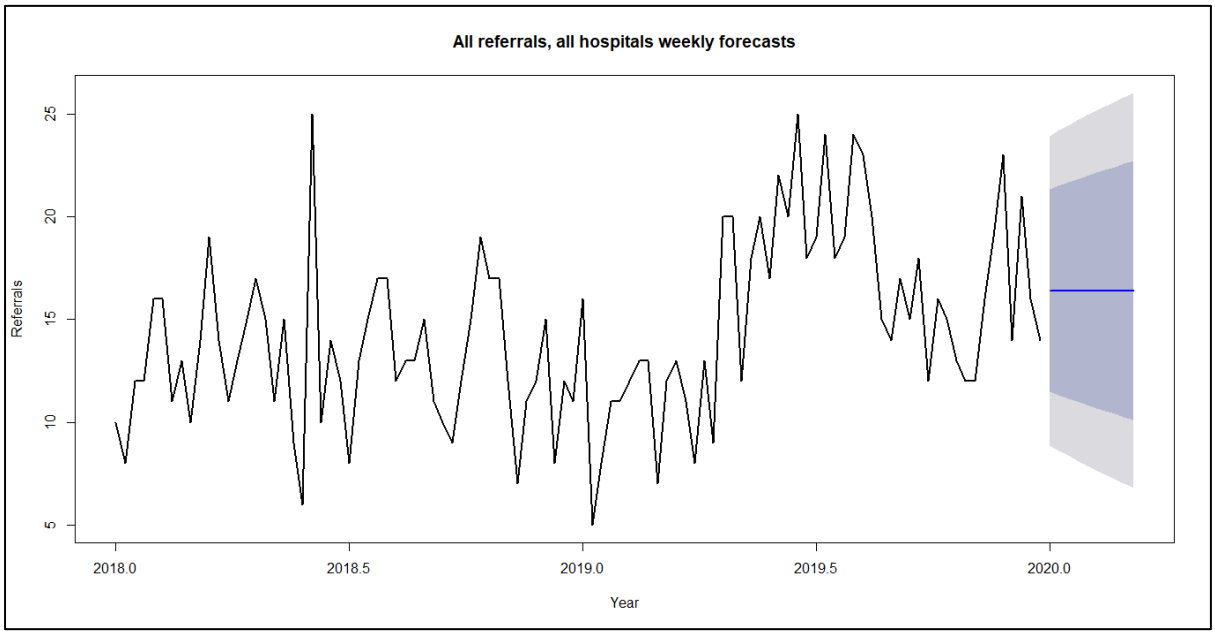

Figure 1.57: forecast- weekly referrals (all hospitals).

Finally, we examine the monthly *referrals* forecasts.

Table 1.25: forecasts from each method- referrals (all hospitals) monthly.

Time series method	Forecast		
Referrals- all hospitals, monthly	July 2019	Aug 2019	Sept 2019
Naïve	89	89	89
Seasonal Naïve	46	68	59
SES	57.76	57.76	57.76
Holt-linear	62	85	63
Holt-Winters	60	74	59
SLR	63.07	63.65	64.25
ARIMA	88.91	87.09	83.26
SSA	57.73	59.44	59.77
ANN	78.84	94.11	95.58

And the error statistics:

Table 1.26: error statistics- monthly referrals (all hospitals) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- all hospitals, monthly training set					
Naïve	11.94	256.81	16.03	-0.05	20.29
Seasonal Naïve	16.20	329.40	18.15	-0.71	26.05
SES	10.81	174.43	13.21	-4.75	18.80
Holt-linear	10.96	166.01	12.88	-4.75	19.45
Holt-Winters	7.42	92.85	9.64	-1.89	12.12
SLR	11	166.09	12.89	-4.69	19.52
ARIMA	14.86	338.16	18.39	-175.93	192.40
SSA	-1.75	122.11	11.05	-1.75	14.40
ANN	2.43	12.60	3.55	-1.23	4.24

Table 1.27: error statistics- monthly referrals (all hospitals) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- all hospitals, monthly test set					
Naïve	16.20	333.80	18.28	14.61	18.20
Seasonal Naïve	18	482.80	21.97	-33.86	33.86
SES	18.24	497.39	22.30	-31.57	31.57
Holt-linear	13.44	321.19	17.92	-18.95	21.14
Holt-Winters	14.08	383.62	19.59	-19.62	22.96
SLR	13.29	315.48	17.76	-18.46	20.83
ARIMA	16.20	333.80	18.27	14.61	18.20
SSA	16.49	435.37	20.87	-27.79	27.79
ANN	15.99	378.44	19.45	14.72	17.54

Here we plot the method with the lowest error statistics from the training set, ANN, and the method with the lowest error statistics from the test set, SLR:

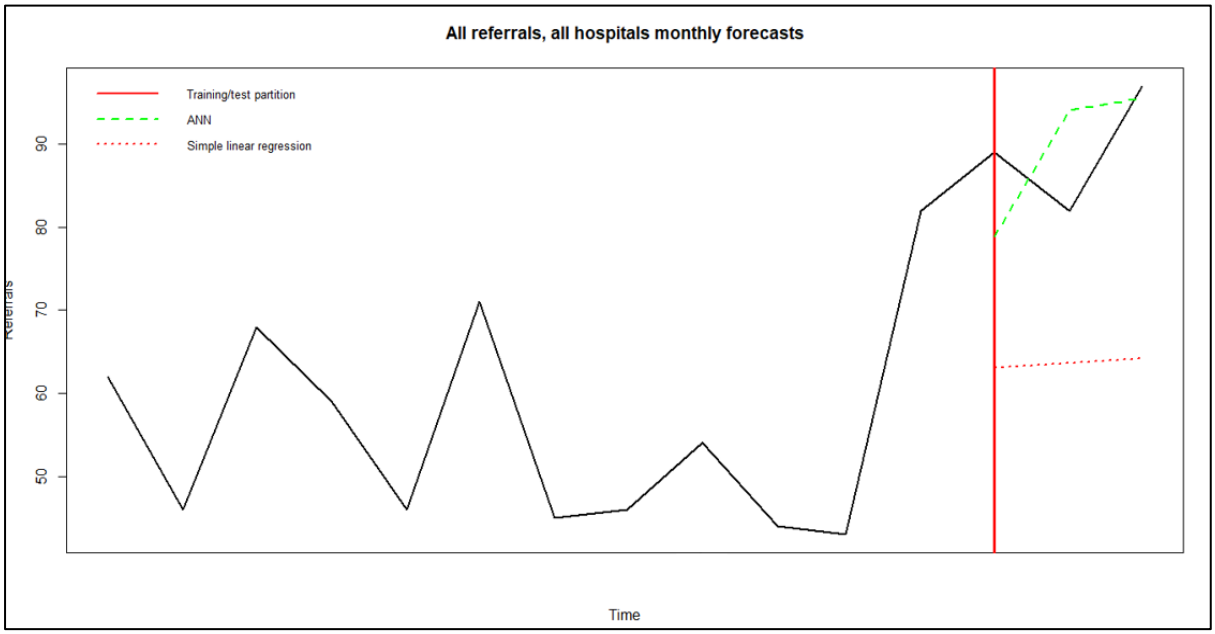

Figure 1.58: forecast- monthly referrals (all hospitals).

Despite its low error statistics, the linear regression model underestimates our data. Given that the ANN again has the lowest error statistics in the training set, a low MPE and MAPE in the test data and reflects the variability in the data, we again endorse this model. We note that, as our time horizon of interest is around 3 months, the ANN fits the data well. Also, the low MPE and MAPE in the training set provide confidence in this model. Our forecast for the next 3 months from this method is given by:

Table 1.28: three-month forecast- monthly referrals (all hospitals).

Month	Forecast (ANN)
Dec 2019	70.31
Jan 2020	80.81
Feb 2020	75.59

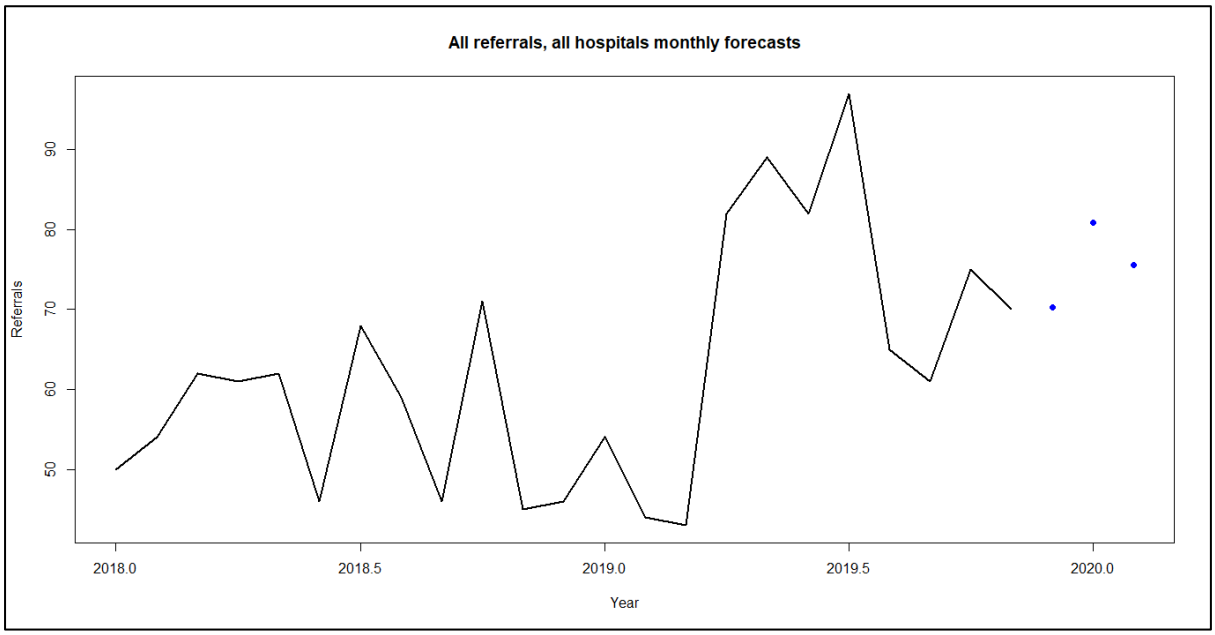

Figure 1.59: forecast- monthly referrals (all hospitals).

We examine the daily/weekly/monthly *tests* data forecasts/results, restricting our attention to all test types at all hospitals.

Table 1.29: forecasts from each method- tests (all types) daily.

Time series method	Forecast (2019)									
Tests- all tests, daily	11/06	12/06	13/06	14/06	15/06	16/06	17/06	18/06	19/06	20/06
Naïve	2	2	2	2	2	2	2	2	2	2
Seasonal Naïve	1	12	3	3	0	1	5	5	14	9
SES	2.34	2.34	2.34	2.34	2.34	2.34	2.34	2.34	2.34	2.34
Holt-linear	2.20	2.19	2.19	2.19	2.19	2.19	2.19	2.19	2.18	2.18
Holt-Winters	7.07	2.41	2.20	0.39	0.20	2.70	2.53	7.08	2.42	2.21
SLR	3.16	3.16	3.16	3.17	3.17	3.17	3.17	3.17	3.17	3.17
MLR	2.39	2.39	2.39	2.39	2.40	2.40	2.40	2.40	2.40	2.40
ARIMA	6.06	2.45	2.82	0.36	0.11	2.96	2.30	5.94	2.44	2.80
SSA	3.52	1.98	2.87	2.80	2.03	3.48	1.63	3.50	1.97	2.85
ANN	7.68	3.70	3.18	0.35	1.37	2.24	3.25	2.51	0	3.37

With corresponding error statistics:

Table 1.30: error statistics- daily tests (all tests) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, daily training set					
Naïve	2.85	18.56	4.31	-Inf	Inf
Seasonal Naïve	3	20.18	4.49	-Inf	Inf
SES	2.48	11.60	3.41	-Inf	Inf
Holt-linear	2.52	11.54	3.40	Inf	Inf
Holt-Winters	1.60	5.56	2.36	Inf	Inf
SLR	2.65	12.21	3.49	-Inf	Inf
MLR	1.60	5.53	2.35	Inf	Inf
ARIMA	1.55	5.35	2.31	Inf	Inf
SSA	2.55	10.60	3.26	-Inf	Inf
ANN	0.01	0	0.01	Inf	Inf

Table 1.31: error statistics- daily tests (all tests) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, daily test set					
Naïve	1.80	8.40	2.90	-30	90
Seasonal Naïve	2.90	13.30	3.65	43.73	65.95
SES	1.94	8.11	2.85	-11.05	82.71
Holt-linear	1.88	8.22	2.87	-18.81	85.79
Holt-Winters	2.84	13.02	3.61	-41.35	133.04
SLR	2.30	8.36	2.89	17.96	72.61
MLR	1.96	8.08	2.84	-8.46	81.66
ARIMA	2.61	11.21	3.35	-41.58	132.47
SSA	2.30	9.90	3.15	-14.76	98.43
ANN	2.77	15.67	3.96	689.62	789.46

As ANN scores the lowest error statistics in the training set and MLR scores the lowest MSE, RMSE, MPE, MAPE statistics, we compare these models:

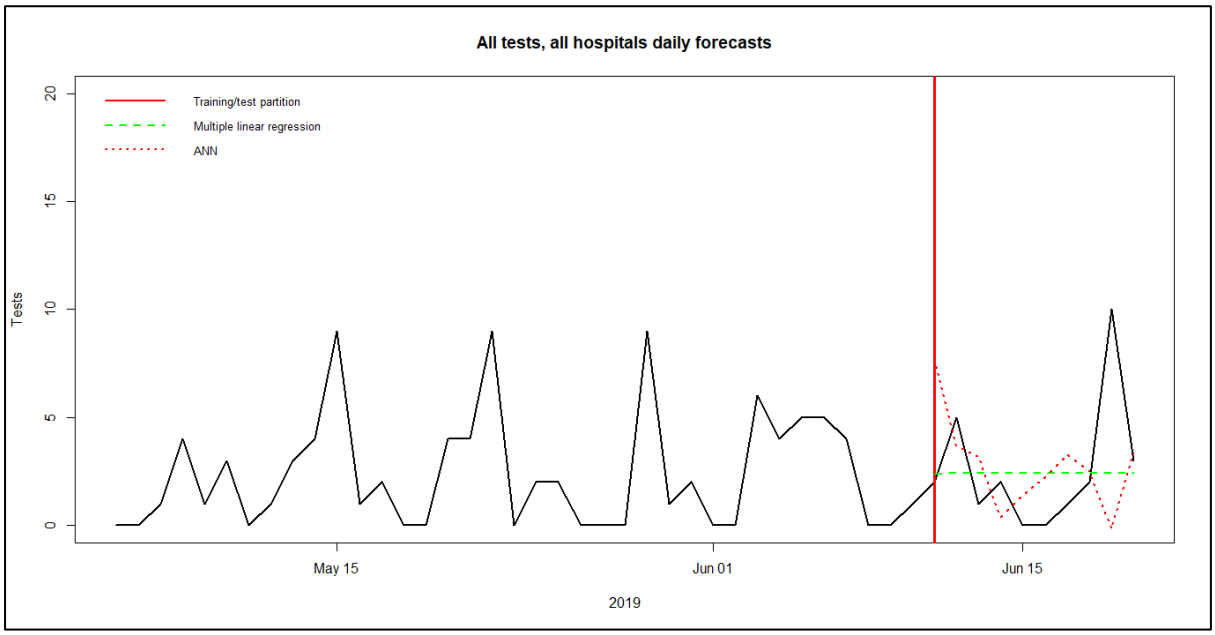

Figure 1.60: forecast- daily tests (all tests).

The MLR model underfits the data. The ANN method, however, overfits the data, and has poor error statistics for the test set. We therefore endorse the MLR method:

Table 1.32: ten-day forecast- daily tests (all tests).

Day	Forecast (MLR)
25/12/2019	0.79
26/12/2019	0.78
27/12/2019	0.77
28/12/2019	0.76
29/12/2019	0.76
30/12/2019	0.75
31/12/2019	0.74
01/01/2020	0.74
02/01/2020	0.73
03/01/2020	0.73

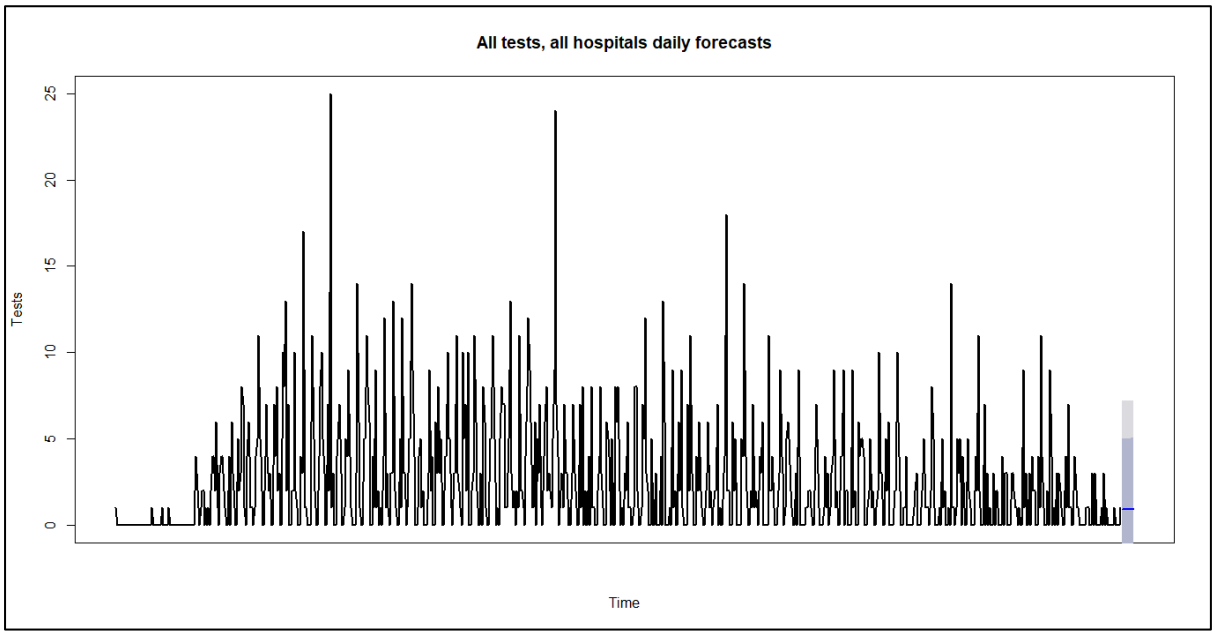

Figure 1.61: forecast- daily tests (all tests).

We expect around 0.75 tests per day. We next assess weekly *tests* data:

Table 1.33: forecasts from each method- tests (all types) weekly.

Time series method	Forecast (2019)									
Tests- all tests, weekly	w/c 06/06	w/c 13/06	w/c 20/06	w/c 27/06	w/c 04/07	w/c 11/07	w/c 18/07	w/c 25/07	w/c 01/08	w/c 08/08
Naïve	17	17	17	17	17	17	17	17	17	17
Seasonal Naïve	31	23	15	23	26	28	28	28	19	28
SES	16.73	16.73	16.73	16.73	16.73	16.73	16.73	16.73	16.73	16.73
Holt-linear	16.19	16.09	15.99	15.88	15.78	15.68	15.58	15.47	15.37	15.27
Holt-Winters	19.43	19	24.53	23.41	18.43	20.44	25.56	26.22	19.42	20.80
SLR	22.05	22.12	22.18	22.24	22.31	22.37	22.43	22.50	22.56	22.62
ARIMA	16.75	16.76	16.76	16.76	16.76	16.76	16.76	16.76	16.76	16.76
SSA	19.89	18.82	17.24	17.03	18.29	19.26	18.47	16.92	16.49	17.51
ANN	20.30	22.52	22.21	12.50	16.85	15.91	17.73	22.61	12.32	14.47

We compare corresponding error statistics:

Table 1.34: error statistics- weekly tests (all tests) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, weekly training set					
Naïve	5.95	61.01	7.81	-Inf	Inf
Seasonal Naïve	10.66	169.72	13.03	-6.39	67.57
SES	4.94	42.53	6.52	-Inf	Inf
Holt-linear	4.89	42.45	6.52	-Inf	Inf
Holt-Winters	4.73	36.63	6.05	-Inf	Inf
SLR	7.35	84.96	9.22	-Inf	Inf
ARIMA	4.92	42.51	6.52	-Inf	Inf
SSA	6.07	65.52	8.09	-Inf	Inf
ANN	0.04	0.01	0.07	-0.04	0.26

Table 1.35: error statistics- weekly tests (all tests) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, weekly test set					
Naïve	4	34.57	5.88	18.82	23.53
Seasonal Naïve	11.90	171.09	13.08	41.52	46.86
SES	3.95	32.95	5.74	17.52	23.59
Holt-linear	3.75	27.35	5.23	12.39	24
Holt-Winters	7.92	87.05	9.33	36.22	36.22
SLR	8.54	97.81	9.89	38.17	38.17
ARIMA	3.95	33.06	5.75	17.65	23.58
SSA	4.86	44.36	6.66	22.84	26.72
ANN	4.94	33.52	5.79	21.72	29

The zeroes in the *tests* dataset result in undefined MPE and MAPE. The ANN method records the lowest error statistics in the training set and the Holt-linear method scores the lowest error statistics in the test set. We compare these:

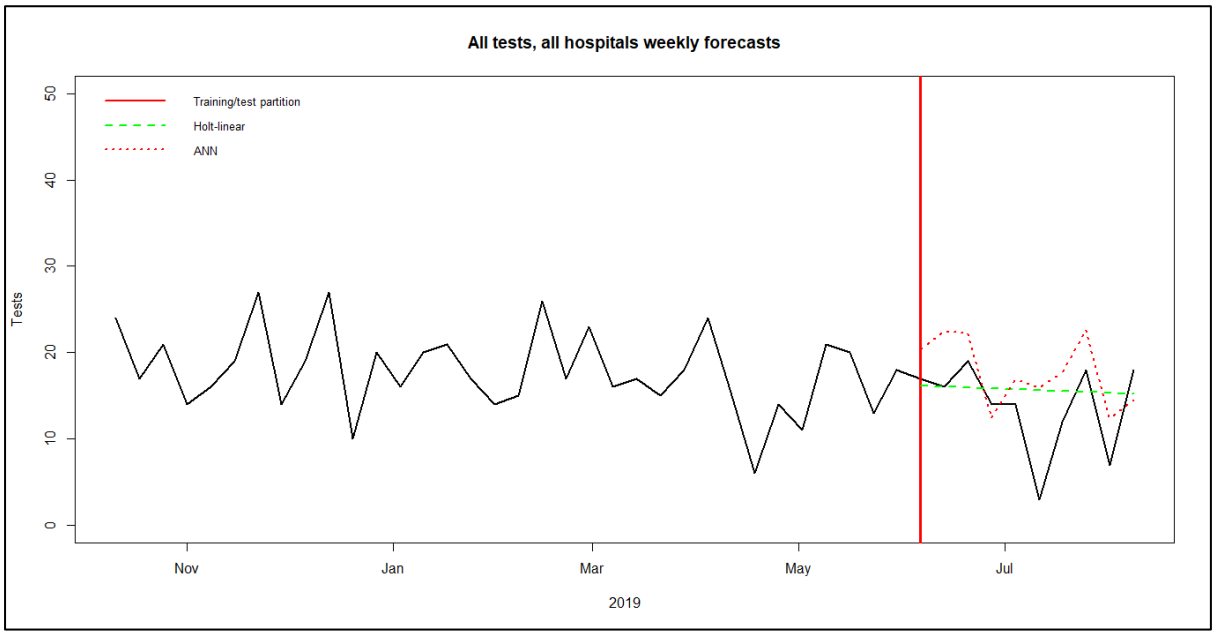

Figure 1.62: forecast- weekly tests (all tests).

The ANN method mostly overestimates the data. The Holt-linear method underfits the data, however its low error statistics in the training and test set provide confidence in this model. This method predicts a continuing downward trend to zero on the entire dataset:

Table 1.36: ten-week forecast- weekly tests (all tests).

Week commencing	Forecast (Holt-Linear)
26/12/2019	2.93
02/01/2020	2.43
09/01/2020	1.94
16/01/2020	1.44
23/01/2020	0.95
30/01/2020	0.45
06/02/2020	0
13/02/2020	0
20/02/2020	0
27/02/2020	0

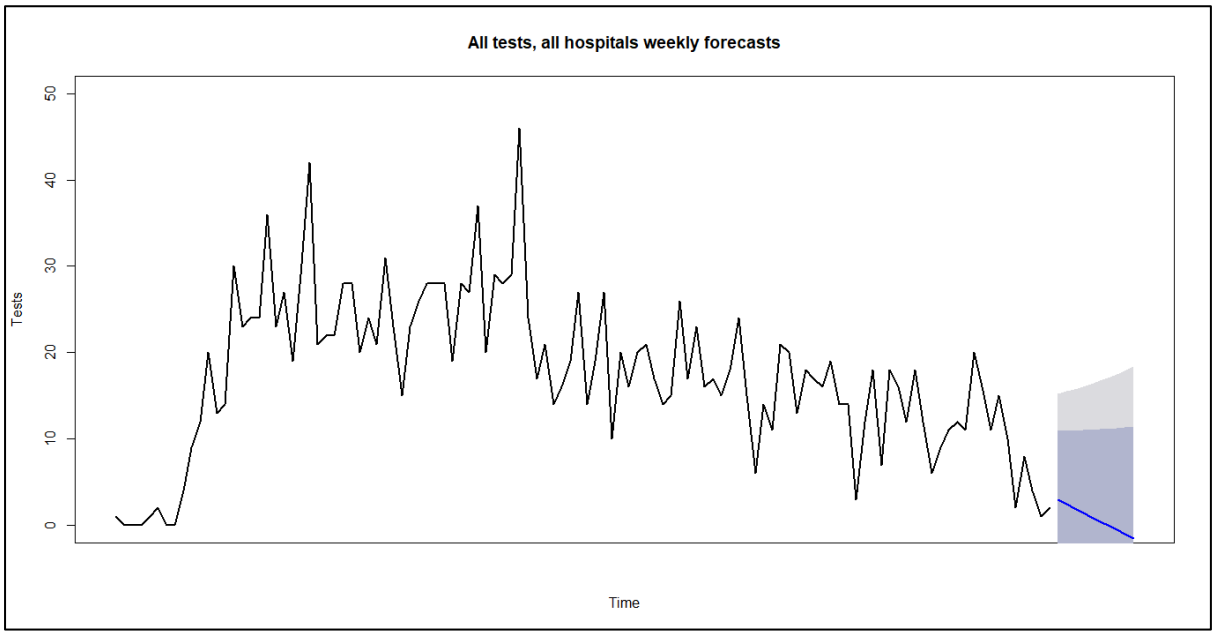

Figure 1.63: forecast- weekly tests (all tests).

Finally, we inspect the monthly *tests* data:

Table 1.37: forecasts from each method- tests (all types) monthly.

Time series method	Forecast		
Tests- all tests, monthly	May 2019	June 2019	July 2019
Naïve	68	68	68
Seasonal Naïve	98	117	117
SES	68	68	68
Holt-linear	70.89	73.78	76.67
Holt-Winters	62.76	60.72	58.67
SLR	97.65	99.02	100.40
ARIMA	88.91	87.09	83.26
SSA	79.52	78.51	76.01
ANN	67.33	67.34	67.26

Table 1.38: error statistics- monthly tests (all tests) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, monthly					
Naïve	15.33	459.44	21.43	4.31	18.79
Seasonal Naïve	46.57	2562	50.62	2.66	62.19
SES	14.53	435.28	20.86	4.10	17.81
Holt-linear	15.70	437.36	20.91	-49.68	68.96
Holt-Winters	21.64	867.95	29.46	1.82	21.64
SLR	28.50	1124.25	33.53	-366.85	392.01
ARIMA	14.86	338.16	18.39	-175.93	192.40
SSA	19.55	784.67	28.01	-396.89	405.32
ANN	0.35	0.21	0.46	-0.02	0.48

Table 1.39: error statistics- monthly tests (all tests) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, monthly test set					
Naïve	10.57	221.12	14.87	15.13	15.55
Seasonal Naïve	45.29	2367.80	48.66	42.78	42.78
SES	10.57	221.12	14.87	15.13	15.55
Holt-linear	21.84	717.70	26.79	26.35	26.35
Holt-Winters*	10.57	221.12	14.87	15.12	15.55
SLR	44.06	2107.73	45.91	43.04	43.04
ARIMA	26.90	815.10	28.55	31.95	31.95
SSA	15.64	320.77	17.91	21.55	21.55
ANN	7.69	94.87	9.74	10.02	11.40

**Holt-Winters could not estimate seasonality due to insufficient data.*

Here the ANN method has the lowest error statistics on both the training and test set. We compare the ANN with another method which scored low error statistics on the test set, SES.

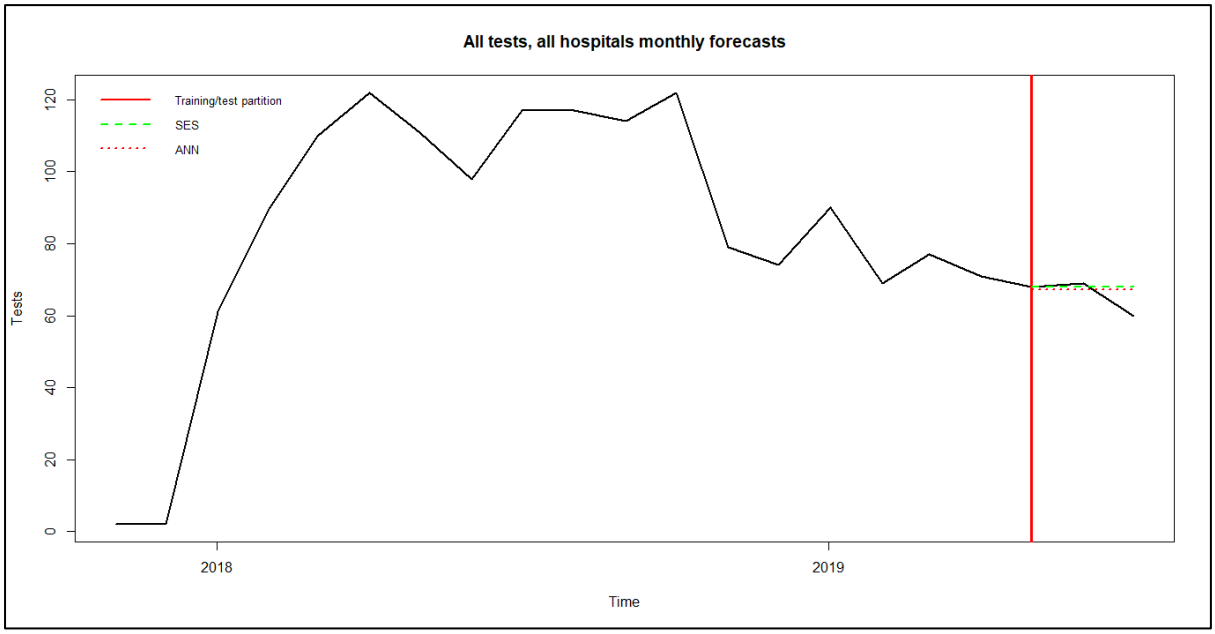

Figure 1.64: forecast- monthly tests (all tests).

Both SES and ANN perform well. As ANN scores the lower test statistic, we endorse the ANN model. We calculate the next 3 months on the entire dataset is:

Table 1.40: three-month forecast- monthly tests (all tests).

Month	Forecast (ANN)
Jan 20	6.48
Feb 20	11.76
Mar 20	8.65

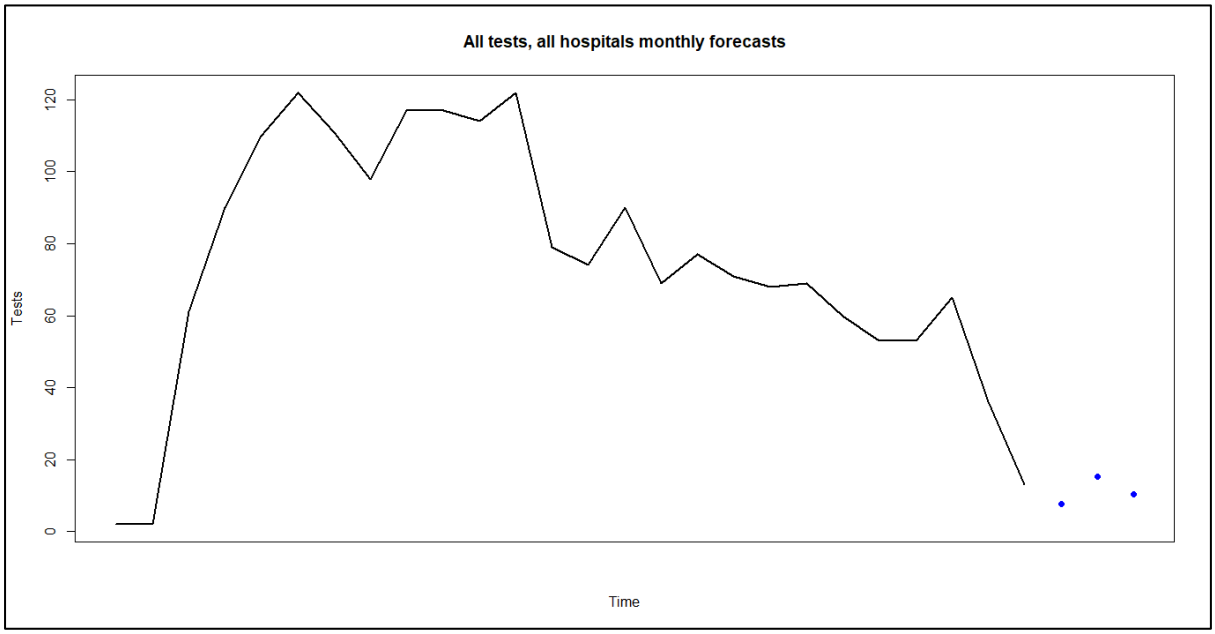

Figure 1.65: forecast- monthly tests (all tests).

1.5.3 Hybrid models

From Zhang (2002), we consider hybrid models, which fit multiple models and ensemble them to create forecasts. We use the *R* command *hybridModel* from package *forecastHybrid* to determine which hybrid model to consider. For brevity we consider the daily/weekly/monthly *referrals* (all hospitals) and daily/weekly/monthly tests data (all tests). The hybrid models we examine are combinations of ARIMA and *theta* methods (we do not discuss theta methods in this paper). The error statistics from such hybrid models are:

Table 1.41: error statistics- hybrid method referrals (all hospitals) training set.

Hybrid method- training set	MAE	MSE	RMSE	MPE	MAPE
Referrals- all tests, daily	1.44	3.13	1.77	-Inf	Inf
Referrals- all tests, weekly	2.95	13.79	3.71	-7.19	25.25
Referrals- all tests, monthly	10.80	174.34	13.20	-4.53	18.75

Table 1.42: error statistics- hybrid method referrals (all hospitals) test set.

Hybrid method- test set	MAE	MSE	RMSE	MPE	MAPE
Referrals- all tests, daily	2.05	4.80	2.19	6.07	86.80
Referrals- all tests, weekly	3.53	17.14	4.14	8.71	16.48
Referrals- all tests, monthly	16.94	453.26	21.29	-28.71	28.71

The daily referrals hybrid model has poor error statistics; we reject this model. The weekly hybrid model performs well compared to other models; we examine a plot of the hybrid model forecast against our favoured ANN model:

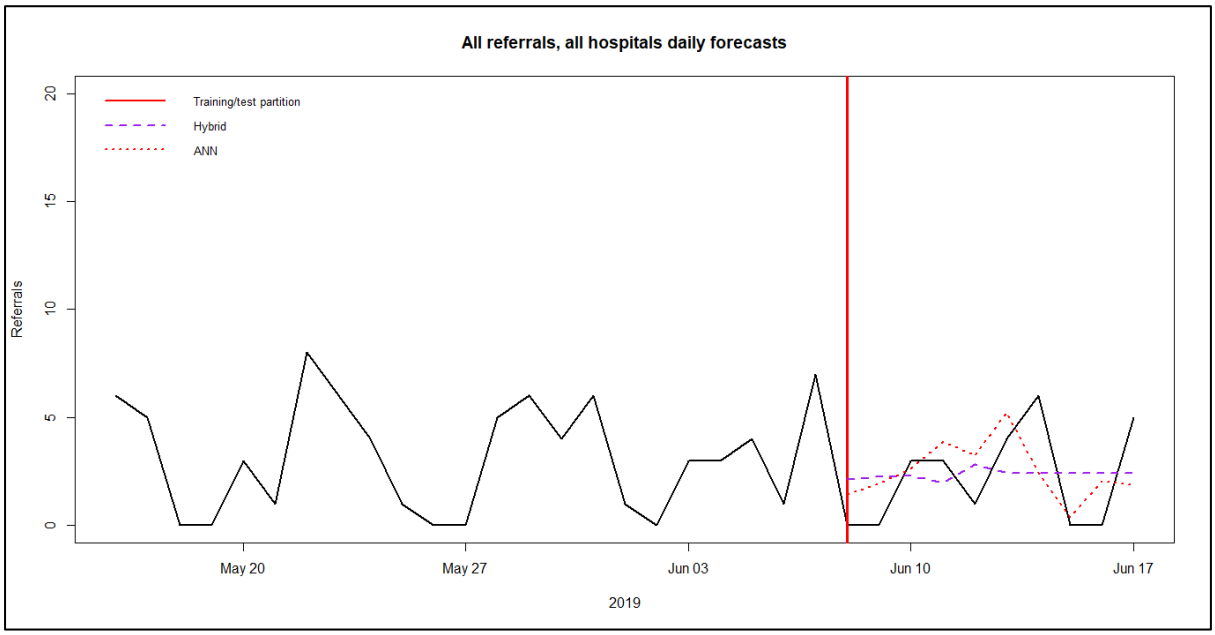

Figure 1.66: forecast- daily referrals (all hospitals).

The hybrid model performs well but does not reflect the noise in the data and has worse error statistics than the ANN model. We still favour the ANN model.

We examine a plot of our favoured ARIMA model against our hybrid model:

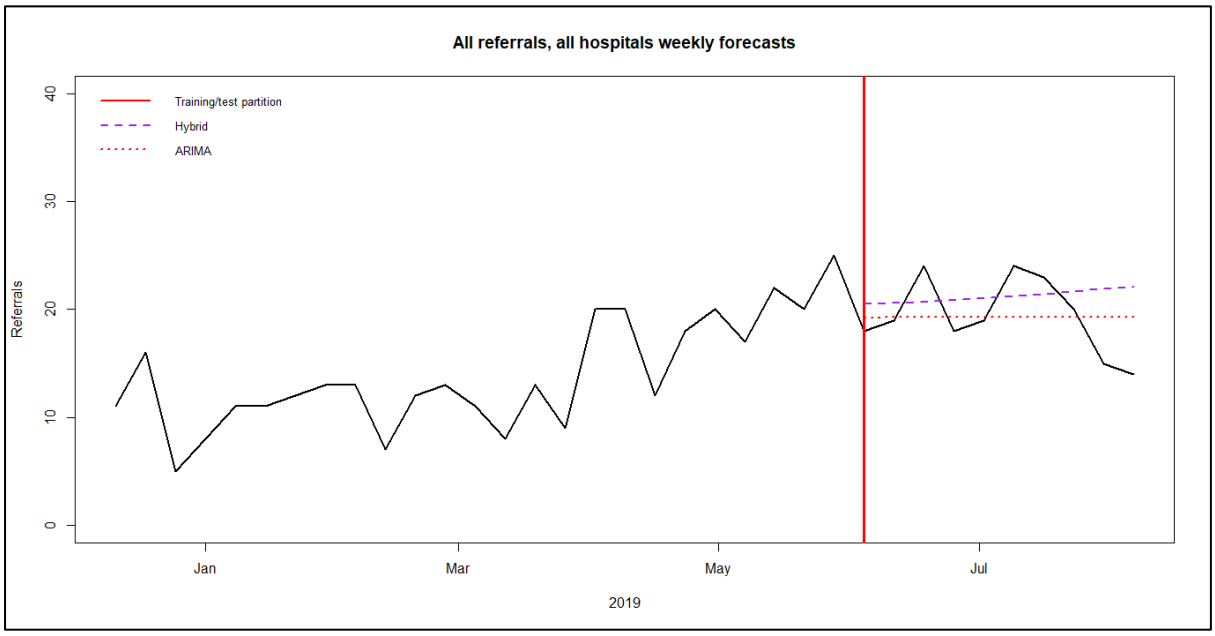

Figure 1.67: forecast- weekly referrals (all hospitals).

Both models appear to fit the data well. As our ARIMA model has lower error statistics than our new hybrid model, we favour the ARIMA model. Finally, we consider monthly *referrals*.

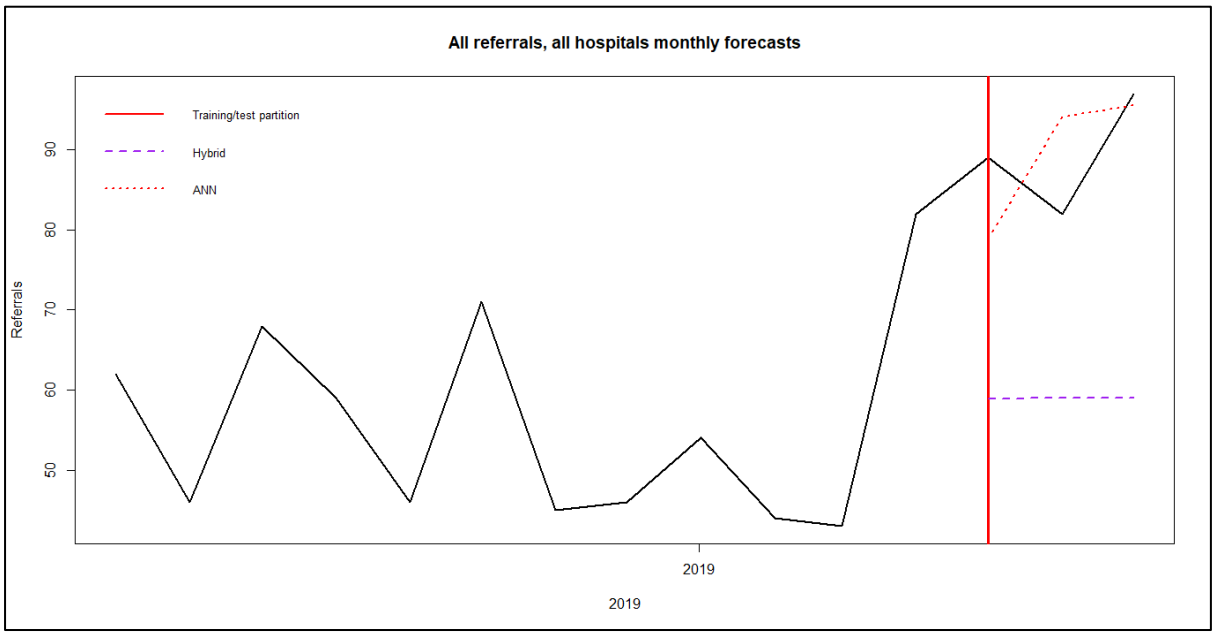

Figure 1.68: forecast- monthly referrals (all hospitals).

The hybrid model clearly offers no improvement on the ANN model for monthly *referrals*.

We consider *tests*:

Table 1.43: error statistics- hybrid method tests (all tests) training set.

Hybrid method- training set	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, daily	2.03	8.54	2.92	Inf	Inf
Tests- all tests, weekly	4.59	37.13	6.09	Inf	Inf
Tests- all tests, monthly	13.91	332.36	18.23	-56.93	74.91

Table 1.44: error statistics- hybrid method tests (all tests) test set.

Hybrid method- test set	MAE	MSE	RMSE	MPE	MAPE
Tests- all tests, daily	1.94	7.34	2.71	7.30	73.90
Tests- all tests, weekly	6.90	81.90	9.05	30.64	31.88
Tests- all tests, monthly	9	165.38	12.86	12.16	13.83

We compare the hybrid method with our favoured MLR on the daily *tests* data:

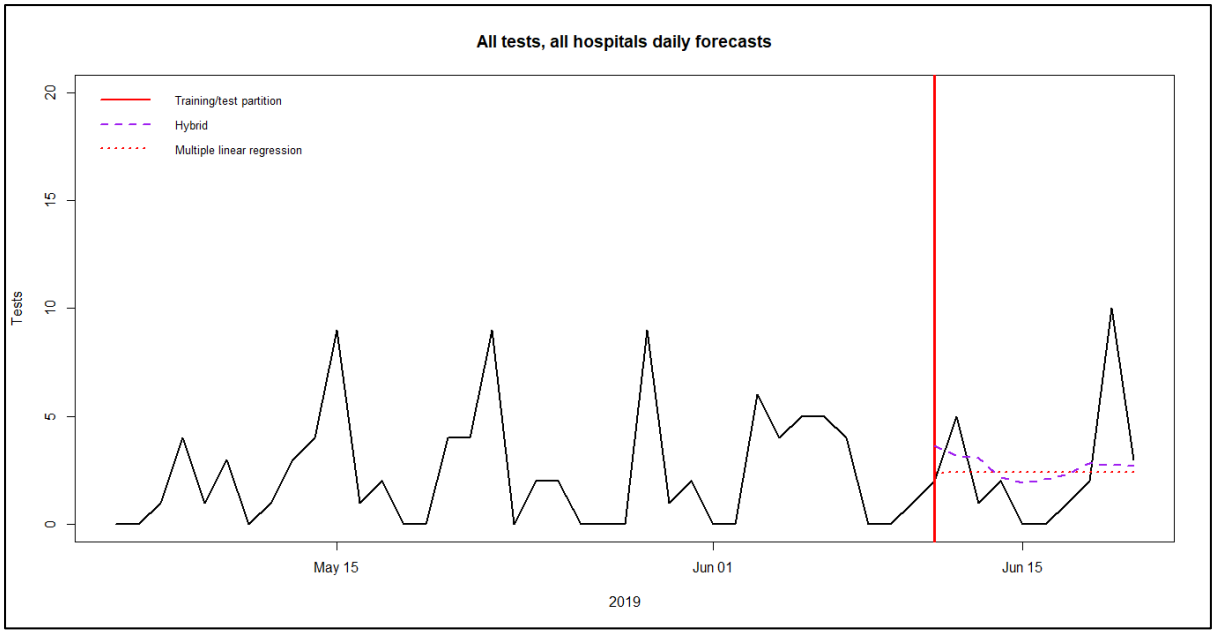

Figure 1.69: forecast- daily tests (all tests).

As the hybrid model does not provide an improved fit, and has larger error statistics, we reject this model, favouring our existing MLR model.

For the weekly data, we compare the forecast from the hybrid method with our favoured ARIMA method:

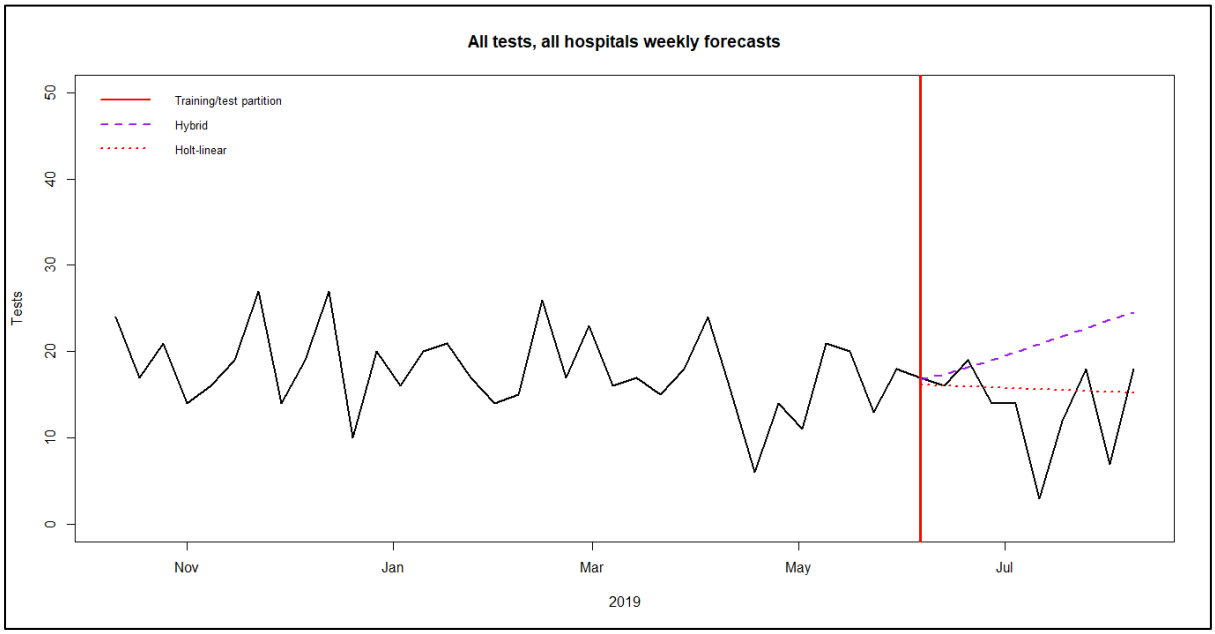

Figure 1.70: forecast- weekly tests (all tests).

Clearly, Holt-linear better models the variation of the data. We therefore keep this model. Finally, we examine the hybrid model on the monthly *tests* data:

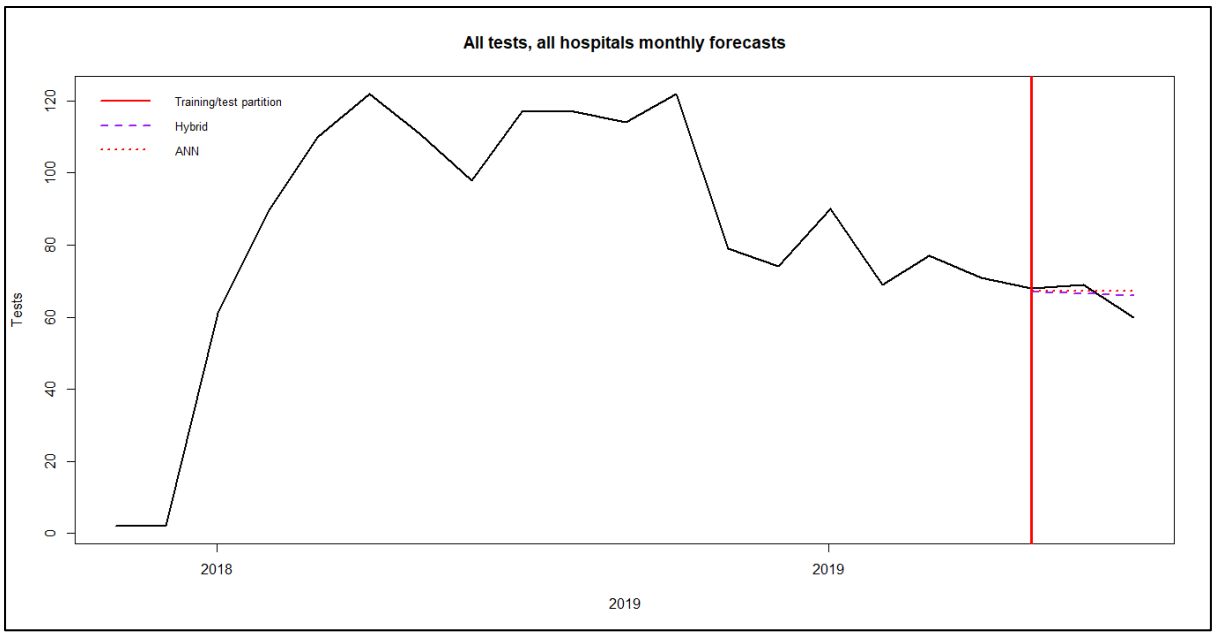

Figure 1.71: forecast- monthly tests (all tests).

As the hybrid method scores larger error statistics than the ANN model, we retain our ANN.

1.5.4 Findings

It is necessary to calculate the error statistics for each method and examine the structure of forecasts produced. Error statistics give evidence about which models provide the best fit and plotting the forecasts from the methods with the lowest error statistics allows us to determine which model to use. The structure of the daily data (with many zeroes), impacts on its MPE and MAPE therefore we should use other metrics to assess the effectiveness of daily models. We should also consider several error statistics before determining which models should be examined further.

For the series of all referrals (daily/weekly/monthly) and all tests (daily/weekly/monthly), the error statistics determine how accurately we can forecast, with interesting results.

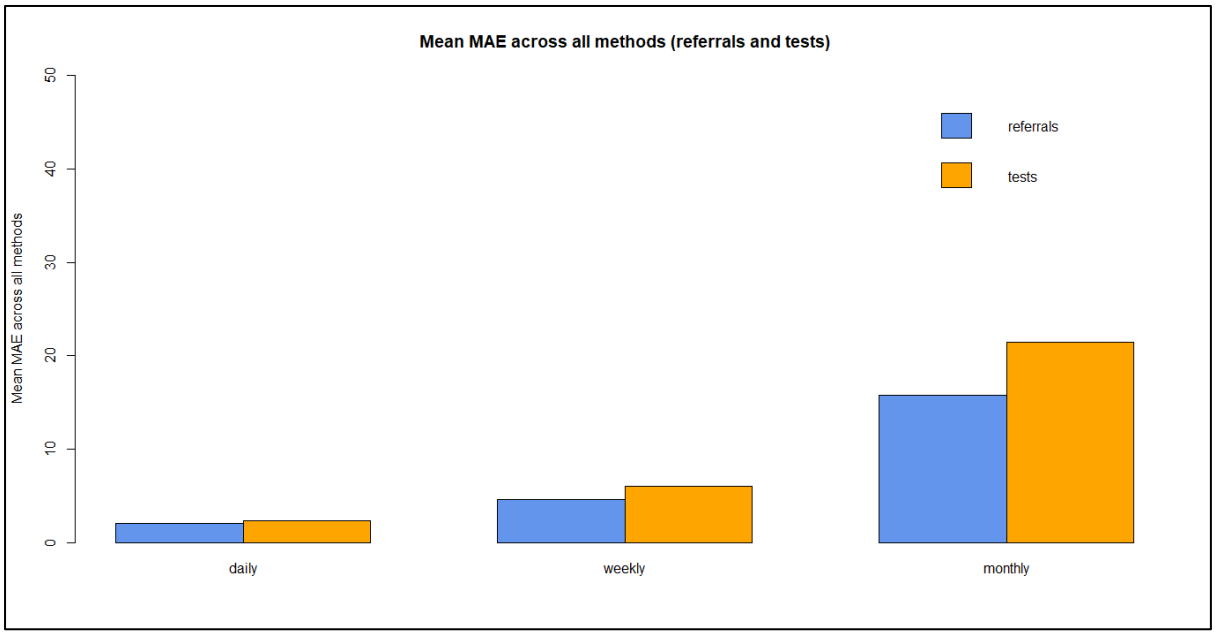

Figure 1.72: mean MAE- all methods (referrals/tests).

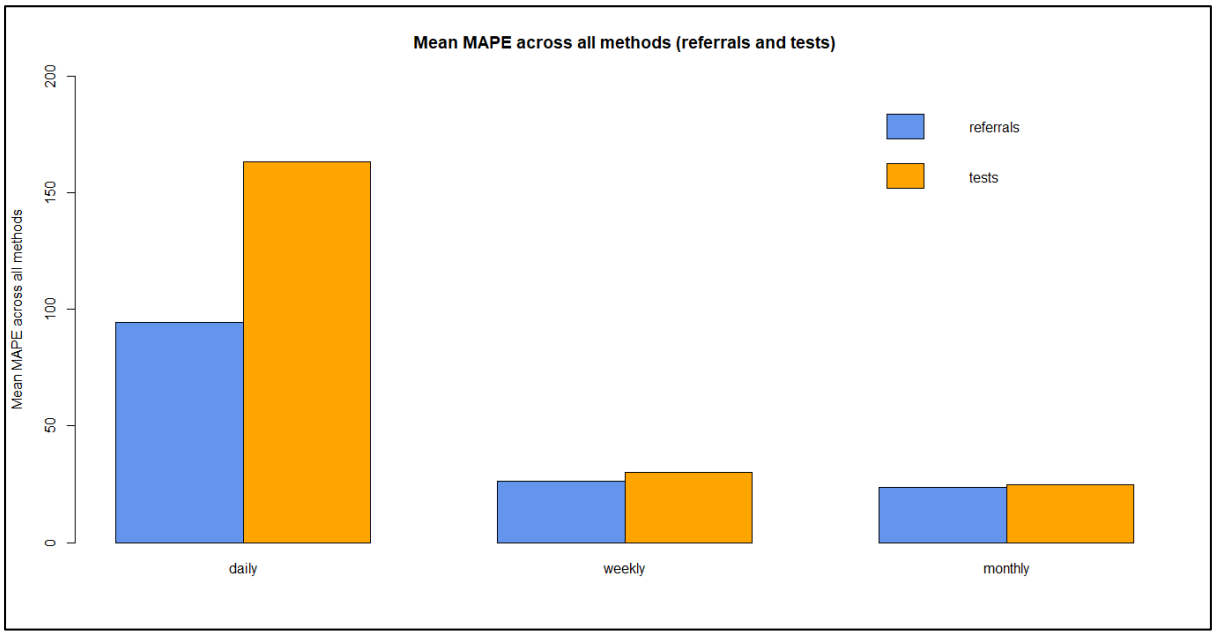

Figure 1.73: mean MAPE- all methods (referrals/tests).

The mean MAE for our daily models is smaller than the weekly/monthly mean MAEs. The MAPE is largest for daily data, and smallest for monthly data. Daily models have the smallest absolute error but the largest absolute percentage error, however the monthly models have the smallest absolute percentage error but the largest absolute error. This is unsurprising, given that there is much more variation in monthly *referrals*/*tests* than the daily data, however this suggests that, when measured as a percentage, our daily forecasts are less accurate than our weekly/monthly forecasts.

The structure of the data affects which model performs best. There is no single model which provides the best fit for every time series examined. Instead, methods which account for trend and/or seasonality (such as ANN, SES and SSA) tend to perform best, whereas more rudimentary naïve models tend to perform worse, particularly in nonstationary data. Hybrid models should be considered for modelling.

Table 1.45: preferred forecasting methods- referrals time series.

Referrals			
	All hospitals	Prince Charles	Royal Glamorgan
Daily	ANN	SSA	ANN
Weekly	ARIMA	ANN	SLR
Monthly	ANN	SES	Holt-Winters

Table 1.46: preferred forecasting methods- tests time series.

Tests				
	All tests	CT tests	PET tests	Other tests
Daily	MLR	-	-	-
Weekly	Holt-linear	SSA	Holt-Winters	SES
Monthly	ANN	ANN	SES	SLR

Table 1.47: frequency table- most preferred forecast method.

Forecasting method	Frequency
ANN	6
SES	3
Holt-Winters	2
SLR	2
SSA	2
ARIMA	1
Holt-linear	1
MLR	1
Total	**18**

ANN models give exceptionally low error statistics when fitted to our time series' training sets. Such models are effective as they generate datapoints learnt from past data, and they have been successfully used in many real-world applications. However, ANNs may be outperformed by extrapolation/simple regression models. ANN models perform well for data with a nonlinear structure[47] such as the daily *referrals* and monthly *tests* data but is outperformed by several linear models for linearly structured data, such as the weekly *referrals* and *weekly tests* data. We should consider the linearity of our data before deciding whether to use ANNs or another robust alternative. As ANNs are prone to overfitting, they are favoured for non-linear, noisy datasets. Linear regression or extrapolation models may be preferred for non-noisy data such as our monthly series.

Despite the weaknesses of ANNs, our experimentation has demonstrated that they are an effective forecasting method. We have recommended the use of an ANN model for 33.33% of the time series discussed here, and it is the most recommended model of any single forecasting method.

1.6 Conclusion

By fitting various time series models to patient administration system data, this report assesses how accurately we can forecast data including lung cancer referrals and lung cancer tests across the CTMUHB. We performed rudimentary analysis on our data to examine stationary and decomposed our time series to assess trend/seasonality/error. We fitted several rudimentary time series methods to forecast future values. We began with naïve models which assume that future values will match past values, and progressed to discuss extrapolation models, which forecast based on some function of past values. After applying these methods to our series, we discussed causal methods (linear regression). Finally, we applied more modern forecasting methods including SSA and ANNs. We extracted the error statistics from each model and discussed which model should be selected for forecasting.

Our time series analysis recommended different forecasting methods for different time series. Consequently, there is no single forecasting method that outperforms all others across all time series and statisticians should consider several different methods before selecting a model.

As an added piece of research, we performed MLR on Welsh cancer datasets to examine links between lung cancer and deprivation levels. Our analysis found significant evidence of a relationship between deprivation and lung cancer incidence. Factors including smoking and historical mining were proposed as explanations for high lung cancer incidence rates in deprived regions of Wales. Further research should examine other factors for high incidences of lung cancer in this region. Research into citizens' lifestyles could provide insight into the causes of high cancer rates.

The scarcity of our data caused difficulty in forecasting accurately which means that forecasting accuracy increases as data increases. With only around two years of data, our extrapolation and machine learning methods were unoptimised. The lack of data also affected the seasonality calculations of our Holt-Winters method; such methods require at least 2 periods of data for meaningful results. Further research should perform the same methods on

longer time series to assess whether comparable results are obtained, and to establish the relationship between time series length and recommended forecasting method.

On non-noisy datasets, ANNs often overfitted the data, therefore statisticians should consider the noise in the time series before recommending this method. For non-noisy datasets, extrapolation models or ARIMA models often outperform ANNs. Further research should assess the performance of other machine learning methods, such as support vector machines or random forest approaches to assess whether these methods are less prone to overfitting. When used correctly on suitable datasets, machine learning methods present an incredibly powerful tool for forecasting. However, such models require much computational power. If we lack such power, extrapolation models and linear regression models are robust alternatives for time series forecasting on non-noisy datasets.

Hybrid models should be considered for time series modelling, though we found little statistical evidence to endorse these models. ANNs performed best on most time series, with extremely low error statistics for the training data, and relatively low statistics for the test data. In several series, the error statistics for ANN and SES/linear regression models were several times smaller than the worst performing method and the naïve method, suggesting that complex methods offer a drastic improvement on rudimentary methods. Our research may be expanded to forecast several months past the time horizon of interest, to determine whether the models we have endorsed are still effective. We should also consider other methods, including different extrapolation models and machine learning techniques, to determine how accurately they forecast short-term, medium-term, and long-term patient administration data.

Time series analysis allows us to robustly and accurately forecast future values. However, relying on past values leaves methods vulnerable to unforeseen extreme circumstances. Crises, which are extreme and unforeseen events, are unpredictable and time series models may not account for extreme occurrences. This is particularly relevant for early 2020, given the extreme and unpredicted COVID-19 pandemic which drastically affected the capacity and operation of standard medical practice worldwide. Our research, however, could prompt further study into the extent to which respiratory clinics were affected by the COVID-19 lockdown.

2. Citations

2.1 References

[1] American Cancer Society, Inc (2019) *Lung Cancer: Canceratlas.cancer.org*, Available at: *https://canceratlas.cancer.org/wp-content/uploads/2019/09/CA3_LungCancer.pdf* (Accessed: 4th July 2020).

[2] Cancer Research UK (2020) *Lung cancer statistics*, Available at: *https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer#heading-Zero* (Accessed: 4th July 2020).

[3] UK Lung Cancer Coalition (2016) *25 by 25: a ten-year strategy to improve lung cancer survival rates*. p. 5.

[4] UK Lung Cancer Coalition (2016). p. 21.

[5] NHS Wales (2014) *: Lung Cancer in Wales, A detailed analysis of population trends of incidence and stage of diagnosis up to and including 2012*, Cardiff, UK. p. 30.

[6] UK Lung Cancer Coalition (2019) *Pathways matter: a review of the implementation of the national lung cancer pathway*. p. 5.

[7] Lung Clinical Expert Group (2017) *National Optimal Lung Cancer Pathway*. p. 4.

[8] NHS Wales (2019) *National optimal pathway for lung cancer: Point of Suspicion to First Definitive Treatment in adults*, Welsh Thoracic Oncology Group (WTOG). p. 2.

[9] Madden, G. & Tan, J. (2007) 'Forecasting telecommunications data with linear models', in *MPRA Paper 14739.* Munich, Germany: University Library of Munich, Germany. p. 34.

[10] Zhang, G. (2003) 'Time series forecasting using a hybrid ARIMA and neural network model' in *Neurocomputing*, 50, pp. 159-175.

[11] Langat, A., Orwa G., Koima, J. (2017) 'Cancer Cases in Kenya; Forecasting Incidents Using Box & Jenkins Arima Model' in *Biomedical Statistics and Informatics*, 2, pp. 37-48.

[12] Ezhil, S. & Vijayalakshmi, C. (2012) 'A Comparative Study in Predicting Colon Rectum Cancer using Auto Regressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) Models' in *International Journal of Computer Applications,* 44(9), pp. 17-22.

[13] Box, G., Tiao, G. (1975) 'Intervention Analysis with Applications to Economic and Environmental Problems' in *Journal of the American Statistical Association,* 70(349), pp. 70-79.

[14] Jani, P. (2014) *Business Statistics: Theory and Applications*, New Delhi, India: PHI Learning Private Limited. p. 432.

[15] R-core@R-project.org, *ts- Time-Series Objects,* Available at: *https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/ts* (Accessed: 7th July 2020).

[16] Gupta, M., Asthana, A., Joshi, N., Mehndiratta, P. (2018) 'Improving time series forecasting using mathematical and deep learning models' in *International Conference on Big Data Analytics, Springer*, pp. 115-125.

[17] Brooks, R., Gray, J. (2004) 'History of the Forecasters' in *Journal of Portfolio Management,* 31(1), pp. 113-117.

[18] Ragsdale, C. (2007) *Spreadsheet Modeling and Decision Analysis: A Practical Introduction to Management Science,* 5 edn., Stamford, CT, USA: South-Western College Pub. p. 486.

[19] Brown, R (1956) *Exponential Smoothing for Predicting Demand,* Available at: *http://legacy.library.ucsf.edu/tid/dae94e00* (Accessed: 12th July 2020).

[20] Dielman, T. (2006) 'Choosing Smoothing Parameters For Exponential Smoothing: Minimizing Sums Of Squared Versus Sums Of Absolute Errors' in *Journal of Modern Applied Statistical Methods,* 5(1), pp. 118-129.

[21] Reid, R., Sanders, N. (2013) *Operations Management: an integrated approach,* 5 edn., Hoboken, NJ, USA: Wiley. p. 287.

[22] Gardner Jr, E. (2006) 'Exponential smoothing: The state of the art—Part II' in *International Journal of Forecasting,* 22, pp. 637-666.

[23] Chatfield, C. (1978) ' The Holt-Winters Forecasting Procedure' in *Journal of the Royal Statistical Society,* 27(3), pp. 264-279.

[24] Chatfield, C., Yar, M. (1988) 'Holt-Winters Forecasting: Some Practical Issues' in *Journal of the Royal Statistical Society,* 37, pp. 129-140.

[25] Goodwin, P. (2010) 'The Holt-Winters Approach to Exponential Smoothing: 50 Years Old and Going Strong' in *Foresight: The International Journal of Applied Forecasting, International Institute of Forecasters,* 19, pp. 30-33.

[26] Hyndman, R. J., & Athanasopoulos, G (2014) *Forecasting: Principles and practice,* Heathmont, Australia: Otexts.com. p. 103.

[27] Jabbar, H., Khan, Dr R. (2015) 'Methods to Avoid Over-Fitting and Under-Fitting in Supervised Machine Learning (Comparative Study)' in J. Stephen, H. Rohil, S. Vasavi (ed.) *Computer Science, Communication & Instrumentation Devices.* pp. 163-172.

[28] Hyndman, R. J., & Athanasopoulos, G (2014), p. 108.

[29] Macmillan Cancer Support (2017) *Deprivation and Survival from Lung Cancer in Scotland*.

[30] Thomas, R. (2018) *Cancer mortality in Wales, 2001-2017,* Available at: *https://phw.nhs.wales/services-and-teams/welsh-cancer-intelligence-and-surveillance-unit-wcisu/cancer-mortality-in-wales-2001-2017/* (Accessed: 20th July 2020).

[31] NHS Wales (2014) p. 4.

[32] NHS Wales: Cwm Taf University Health Board (2015) *Annual Report of the Director of Public Health 2015*, NHS Wales. p. 4.

[33] Prabhakaran, S. *ARIMA Model – Complete Guide to Time Series Forecasting in Python,* Available at: *https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/* (Accessed: 25th July 2020).

[34] Hyndman, R. J., & Athanasopoulos, G (2014), p. 74.

[35] McLeod, A. (1993) 'Parsimony, model adequacy and periodic correlation in time series forecasting' in *International Statistical Review*, 61, 387-393.

[36] Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. A. (2001). *Analysis of time series structure: SSA and related techniques*, Boca Raton, FL, USA, Chapman and Hall. p. 1.

[37] Zhigljavsky A. (2011) 'Singular Spectrum Analysis for Time Series' in Lovric M. (eds) *International Encyclopedia of Statistical Science*, 3, Springer.

[38] Bojang, P.O., Yang, T.-C., Pham, Q.B., Yu, P.-S. (2020) 'Linking Singular Spectrum Analysis and Machine Learning for Monthly Rainfall Forecasting' in *Appl. Sci.,* 10(9).

[39] D'arcy, J. *Decomposing Time Series Data With Singular-Spectrum Analysis,* Available at: *https://www.kaggle.com/jdarcy/introducing-ssa-for-time-series-decomposition* (Accessed: 30th July 2020).

[40] Gershenson, C. *Artificial Neural Networks for Beginners,* Available at: *https://arxiv.org/ftp/cs/papers/0308/0308031.pdf* (Accessed: 8th August 2020).

[41] Beck, M. *Visualizing neural networks in R,* Available at: *https://beckmw.wordpress.com/2013/11/14/visualizing-neural-networks-in-r-update/* (Accessed: 8th August 2020).

[42] Zhang, G., Qi, M. (2005) 'Neural network forecasting for seasonal and trend time series' in *European Journal of Operational Research* , 160, pp. 501-514.

[43] Zhou, T. (2020) *Deep Learning for Time Series and why deep learning?,* Available at: *https://towardsdatascience.com/deep-learning-for-time-series-and-why-deep-learning-a6120b147d60* (Accessed: 12th August 2020).

[44] Tu, J. (1996) 'Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes' in *Journal of Clinical Epidemiology,* 49(11), pp. 1225-1231.

[45] Pascual, C. (2018) *Tutorial: Understanding Regression Error Metrics in Python,* Available at: *https://www.dataquest.io/blog/understanding-regression-error-metrics* (Accessed: 15th August 2020).

[46] Makridakis S., Spiliotis E., Assimakopoulos V. (2018) *Statistical and Machine Learning forecasting methods: Concerns and ways forward*, 3 edn., Universidad Veracruzana, Mexico: PloS ONE.

[47] Adebiyi, A., Adewumi, A., Ayo, C. (2014) 'Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction' in *Journal of Applied Mathematics,* 1, pp. 1-7.

2.2 Bibliography

Adebiyi, A., Adewumi A., Ayo, C. (2014) 'Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction' in *Journal of Applied Mathematics*, 1, pp. 1-7.

American Cancer Society, Inc (2019) *Lung Cancer: Canceratlas.cancer.org*, Available at: *https://canceratlas.cancer.org/wp-content/uploads/2019/09/CA3_LungCancer.pdf* (Accessed: 4th July 2020).

Beck, M., *Visualizing neural networks in R,* Available at: *https://beckmw.wordpress.com/2013/11/14/visualizing-neural-networks-in-r-update/* (Accessed: 8th August 2020).

Bojang, P.O.; Yang, T.-C.; Pham, Q.B.; Yu, P.-S. (2020) 'Linking Singular Spectrum Analysis and Machine Learning for Monthly Rainfall Forecasting' in *Appl. Sci*., 10(9).

Box, G., Tiao G. (1975) 'Intervention Analysis with Applications to Economic and Environmental Problems' in *Journal of the American Statistical Association*, 70(349), pp. 70-79.

Brooks, R., Gray, J.B. (2004) 'History of the Forecasters' in *Journal of Portfolio Management*, 31(1), pp. 113-117.

Brown, R.G (1956) *Exponential Smoothing for Predicting Demand*, Available at: *http://legacy.library.ucsf.edu/tid/dae94e00* (Accessed: 12th July 2020).

Cancer Research UK (2020) Lung cancer statistics, Available at: *https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer#heading-Zero* (Accessed: 4th July 2020).

Chatfield, C. (1978) ' The Holt-Winters Forecasting Procedure' in *Journal of the Royal Statistical Society*, 27(3), pp. 264-279.

Chatfield, C., Yar, M. (1988) 'Holt-Winters Forecasting: Some Practical Issues' in *Journal of the Royal Statistical Society*, 37, pp. 129-140.

D'arcy, J., *Decomposing Time Series Data With Singular-Spectrum Analysis*, Available at: *https://www.kaggle.com/jdarcy/introducing-ssa-for-time-series-decomposition* (Accessed: 30th July 2020).

Dielman, T. (2006) 'Choosing Smoothing Parameters For Exponential Smoothing: Minimizing Sums Of Squared Versus Sums Of Absolute Errors' in *Journal of Modern Applied Statistical Methods*, 5(1), pp. 118-129.

Ezhil, S, Vijayalakshmi, C. (2012) 'A Comparative Study in Predicting Colon Rectum Cancer using Auto Regressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) Models' in *International Journal of Computer Applications*, 44(9), pp. 17-22.

Gardner Jr., E. (2006) ‘Exponential smoothing: The state of the art—Part II’ in *International Journal of Forecasting*, 22, pp. 637-666.

Gershenson, C., *Artificial Neural Networks for Beginners*, Available at: *https://arxiv.org/ftp/cs/papers/0308/0308031.pdf* (Accessed: 8th August 2020).

Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. A. (2001). *Analysis of time series structure: SSA and related techniques*, Boca Raton, FL, USA, Chapman and Hall.

Goodwin, P. (2010) ‘The Holt-Winters Approach to Exponential Smoothing: 50 Years Old and Going Strong’ in *Foresight: The International Journal of Applied Forecasting*, International Institute of Forecasters, 19, pp. 30-33.

Gupta, M., Asthana, A., Joshi, N., Mehndiratta, P. (2018) ‘Improving time series forecasting using mathematical and deep learning models’ in *International Conference on Big Data Analytics*, Springer, pp. 115-125.

Hyndman, R. J., & Athanasopoulos, G (2014) *Forecasting: Principles and practice*, Heathmont, Australia: Otexts.com.

Jabbar, H. Khan, Dr. R. (2015) ‘Methods to Avoid Over-Fitting and Under-Fitting in Supervised Machine Learning (Comparative Study)’ in J. Stephen, H. Rohil, S. Vasavi (ed.) *Computer Science, Communication & Instrumentation Devices*. pp. 163-172.

Jani, P. (2014) *Business Statistics: Theory and Applications*, New Delhi, India: PHI Learning Private Limited.

Langat, A., Orwa, G., Koima, J. (2017) ‘Cancer Cases in Kenya; Forecasting Incidents Using Box & Jenkins Arima Model’ in *Biomedical Statistics and Informatics*, 2, pp. 37-48.

Lung Clinical Expert Group (2017) *National Optimal Lung Cancer Pathway*.

Macmillan Cancer Support (2017) *Deprivation and Survival from Lung Cancer in Scotland*.

Madden, Gary & Tan, Joachim (2007) 'Forecasting telecommunications data with linear models', in *MPRA Paper 14739*. Munich, Germany: University Library of Munich, Germany.

Makridakis, S., Spiliotis, E., Assimakopoulos, V. (2018) *Statistical and Machine Learning forecasting methods: Concerns and ways forward*, 3 edn., Universidad Veracruzana, Mexico: PloS ONE.

McLeod, A. (1993) 'Parsimony, model adequacy and periodic correlation in time series forecasting' in *International Statistical Review*, 61, 387-393.

NHS Wales (2014) : *Lung Cancer in Wales, A detailed analysis of population trends of incidence and stage of diagnosis up to and including 2012*, Cardiff, UK.

NHS Wales (2019) *National optimal pathway for lung cancer: Point of Suspicion to First Definitive Treatment in adults*, Welsh Thoracic Oncology Group (WTOG)..

NHS Wales: Cwm Taf University Health Board (2015) *Annual Report of the Director of Public Health 2015*, NHS Wales.

Pascual, C. (2018) *Tutorial: Understanding Regression Error Metrics in Python*, Available at: *https://www.dataquest.io/blog/understanding-regression-error-metrics* (Accessed: 15th August 2020).

Prabhakaran, S., *ARIMA Model – Complete Guide to Time Series Forecasting in Python*, Available at: *https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/* (Accessed: 25th July 2020).

R-core@R-project.org, *ts- Time-Series Objects*, Available at: *https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/ts* (Accessed: 7th July 2020).

Ragsdale, C. (2007) *Spreadsheet Modeling and Decision Analysis: A Practical Introduction to Management Science*, 5 edn., Stamford, CT, USA: South-Western College Pub.

Reid, R., Sanders, N. (2013) *Operations Management: an integrated approach*, 5 edn., Hoboken, NJ, USA: Wiley.

Thomas, R. (2018) *Cancer mortality in Wales, 2001-2017*, Available at: *https://phw.nhs.wales/services-and-teams/welsh-cancer-intelligence-and-surveillance-unit-wcisu/cancer-mortality-in-wales-2001-2017/* (Accessed: 20th July 2020).

Tu, J. (1996) 'Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes' in *Journal of Clinical Epidemiology*, 49(11), pp. 1225-1231.

UK Lung Cancer Coalition (2016) *25 by 25: a ten-year strategy to improve lung cancer survival rates.*

UK Lung Cancer Coalition (2019) *Pathways matter: a review of the implementation of the national lung cancer pathway*.

Zhang, G. (2003) 'Time series forecasting using a hybrid ARIMA and neural network model', in *Neurocomputing*, 50, pp. 159-175.

Zhang, G. Qi, M., (2005) 'Neural network forecasting for seasonal and trend time series' in *European Journal of Operational Research*, 160, pp. 501-514.

Zhigljavsky A. (2011) 'Singular Spectrum Analysis for Time Series' in Lovric M. (eds) *International Encyclopedia of Statistical Science*, 3, Springer.

Zhou, T. (2020) *Deep Learning for Time Series and why deep learning?*, Available at: *https://towardsdatascience.com/deep-learning-for-time-series-and-why-deep-learning-a6120b147d60* (Accessed: 12th August 2020).

3. Appendix

3.1 Results: all referrals, Prince Charles daily

Analysis- Prince Charles referrals, daily

Table 3.1: forecasts from each method- referrals (PCH), daily.

Time series method	Forecast (2019)									
Referrals- Prince Charles, daily	08/06	09/06	10/06	11/06	12/06	13/06	14/06	15/06	16/06	17/06
Naïve	0	0	0	0	0	0	0	0	0	0
Seasonal Naïve	0	3	0	0	0	0	0	0	0	3
SES	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Holt-linear	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.15	1.15	1.15
Holt-Winters	0.30	1.34	1.55	1.76	1.89	1.57	0.47	0.32	1.36	1.57
Simple linear regression	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
Multiple linear regression	1	1	1	1	1	1	1	1	1	1
ARIMA	0.51	1.13	0.96	0.74	1.17	2.34	0.73	0.84	0.91	0.81
SSA	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.83
ANN	0	0	0.29	0	1.79	1.13	0	1.05	0.59	0.79

Table 3.2: error statistics- daily referrals (PCH) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Prince Charles, daily training set					
Naïve	1.12	2.63	1.62	-Inf	Inf
Seasonal Naïve	1.05	2.39	1.55	-Inf	Inf
SES	0.94	1.44	1.20	-Inf	Inf
Holt-linear	0.98	1.49	1.22	-Inf	Inf
Holt-Winters	0.81	1.17	1.08	Inf	Inf
Simple linear regression	0.95	1.44	1.20	-Inf	Inf
Multiple linear regression	0.79	1.10	1.05	Inf	Inf
ARIMA	0.86	1.22	1.11	Inf	Inf
SSA	0.95	1.41	1.19	-Inf	Inf
ANN	0.02	0	0.03	Inf	Inf

Table 3.3: error statistics- daily referrals (PCH) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Prince Charles, daily test set					
Naïve	0.80	1.60	1.26	-Inf	Inf
Seasonal Naïve	1	2.20	1.48	-Inf	Inf
SES	0.80	1	1	19.23	80.77
Holt-linear	0.88	1.07	1.04	30.06	77.37
Holt-Winters	1.02	1.49	1.22	5.47	118.30
Simple linear regression	0.80	0.98	0.99	13.25	86.75
Multiple linear regression	0.80	1	1	20.17	79.96
ARIMA	0.92	1.20	1.10	11.05	103.35
SSA	0.80	0.96	0.98	4.31	95.69
ANN	1.13	2.13	1.46	-200.59	314.30

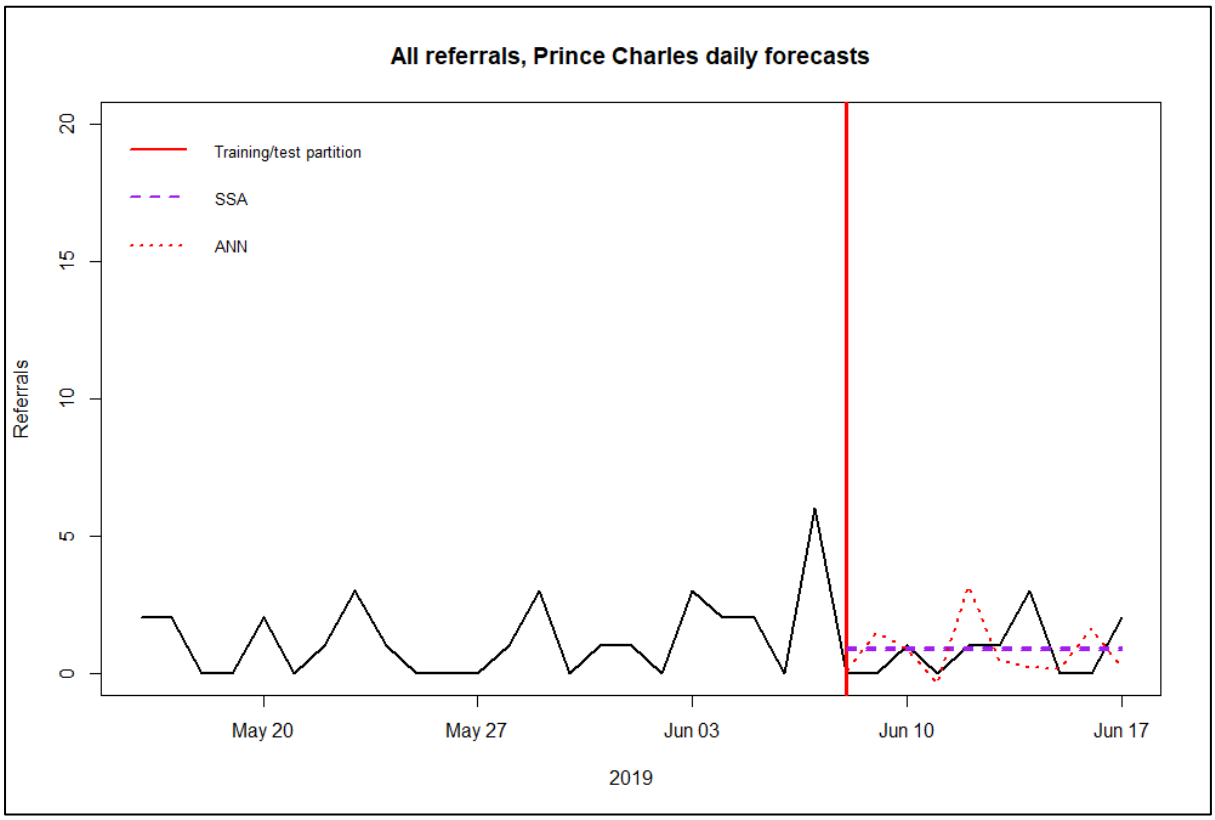

Figure 3.1: forecast- daily referrals (PCH).

We endorse SSA method.

Table 3.4: ten-day forecast- daily referrals (PCH).

Date	Forecast (SSA)
01/12/2019	0.80
02/12/2019	0.85
03/12/2019	1
04/12/2019	0.78
05/12/2019	0.86
06/12/2019	0.99
07/12/2019	0.78
08/12/2019	0.88
09/12/2019	0.99
10/12/2019	0.76

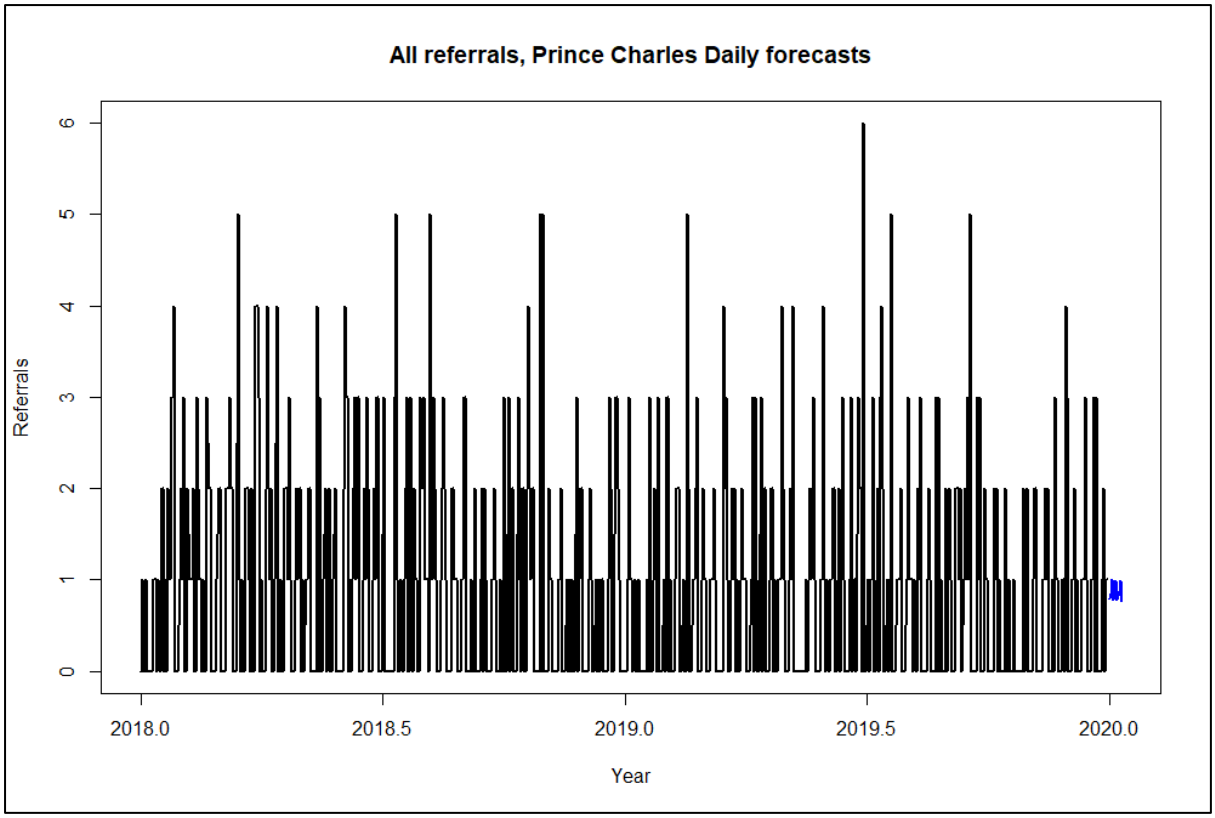

Figure 3.2: forecast- daily referrals (PCH).

3.2 Results: all referrals, Royal Glamorgan daily

Analysis- Royal Glamorgan referrals, daily

Table 3.5: forecasts from each method- referrals (RGH), daily.

Time series method	Forecast (2019)									
Referrals- Royal Glamorgan, daily	08/06	09/06	10/06	11/06	12/06	13/06	14/06	15/06	16/06	17/06
Naïve	0	0	0	0	0	0	0	0	0	0
Seasonal Naïve	0	1	2	1	0	0	0	1	1	0
SES	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
Holt-linear	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74
Holt-Winters	0.10	1.13	0.99	1.15	0.93	1.13	0	0.09	1.13	0.99
Simple linear regression	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74
Multiple linear regression	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
ARIMA	0.78	0.88	0.85	0.83	0.87	0.80	0.78	0.78	0.79	0.79
SSA	0.81	0.80	0.77	0.74	0.71	0.69	0.68	0.68	0.70	0.73
ANN	0	0	0	1.97	1.11	1.07	0.51	0.52	0.24	0.67

Table 3.6: error statistics- daily referrals (RGH) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Royal Glamorgan, daily training set					
Naïve	0.94	1.82	1.35	-Inf	Inf
Seasonal Naïve	0.91	1.89	1.38	-Inf	Inf
SES	0.83	1.07	1.03	-Inf	Inf
Holt-linear	0.83	1.10	1.05	-Inf	Inf
Holt-Winters	0.66	0.89	0.94	Inf	Inf
Simple linear regression	0.82	1.07	1.03	-Inf	Inf
Multiple linear regression	0.65	0.83	0.91	Inf	Inf
ARIMA	0.79	1.02	1	Inf	Inf
SSA	0.81	1.04	1.02	-Inf	Inf
ANN	0	0	0	Inf	Inf

Table 3.7: error statistics- daily referrals (RGH) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Royal Glamorgan, daily test set					
Naïve	0.40	0.60	0.77	-Inf	Inf
Seasonal Naïve	0.80	1.20	1.10	-Inf	Inf
SES	0.72	0.60	0.77	49.95	90.05
Holt-linear	0.69	0.55	0.74	45.73	94.27
Holt-Winters	0.74	0.72	0.85	1587.39	1602.66
Simple linear regression	0.70	0.55	0.74	45.87	94.13
Multiple linear regression	0.68	0.53	0.73	42.63	97.37
ARIMA	0.73	0.62	0.79	50.11	89.89
SSA	0.71	0.58	0.76	42.75	97.25
ANN	0.51	0.46	0.68	456	525.68

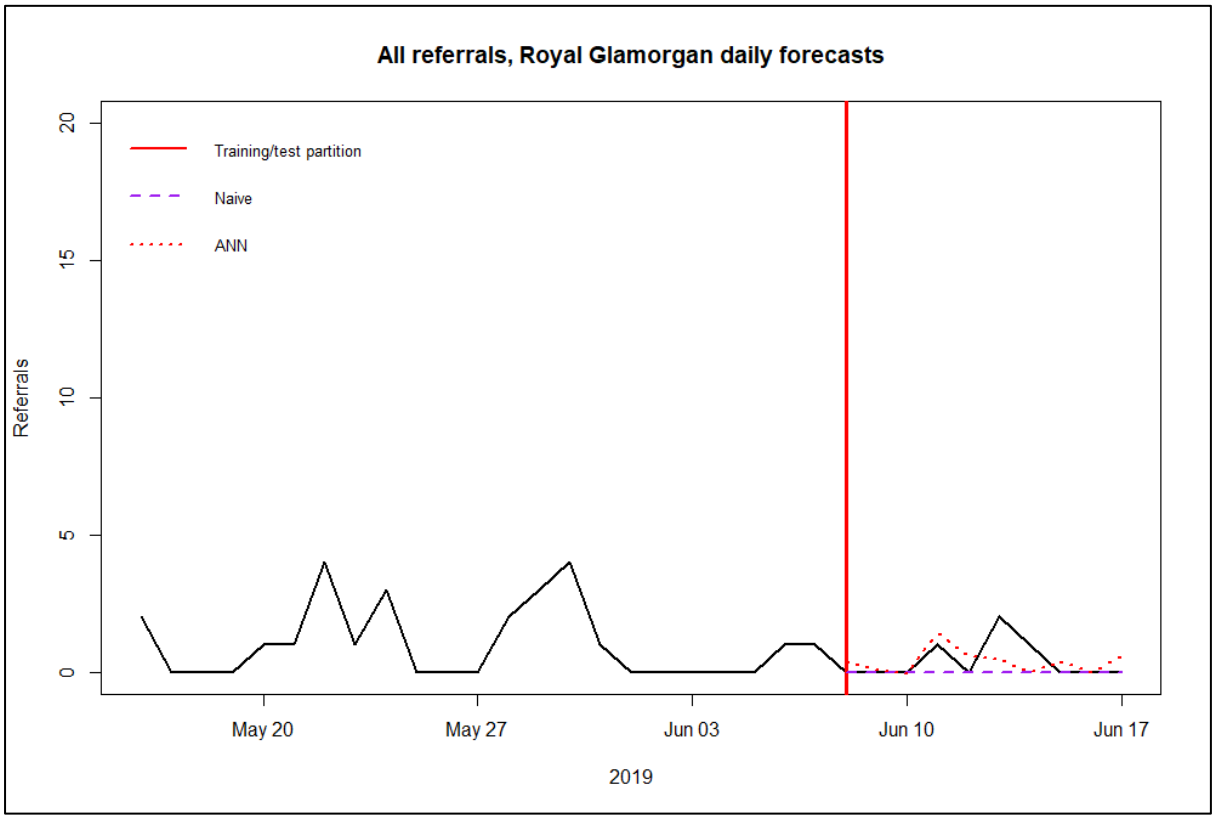

Figure 3.3: forecast- daily referrals (RGH).

We endorse ANN method.

Table 3.8: ten-day forecast- daily referrals (RGH).

Date	Forecast (ANN)
01/12/2019	0
02/12/2019	0
03/12/2019	1.24
04/12/2019	0.81
05/12/2019	0.90
06/12/2019	1.11
07/12/2019	0
08/12/2019	0.59
09/12/2019	1.73
10/12/2019	0.98

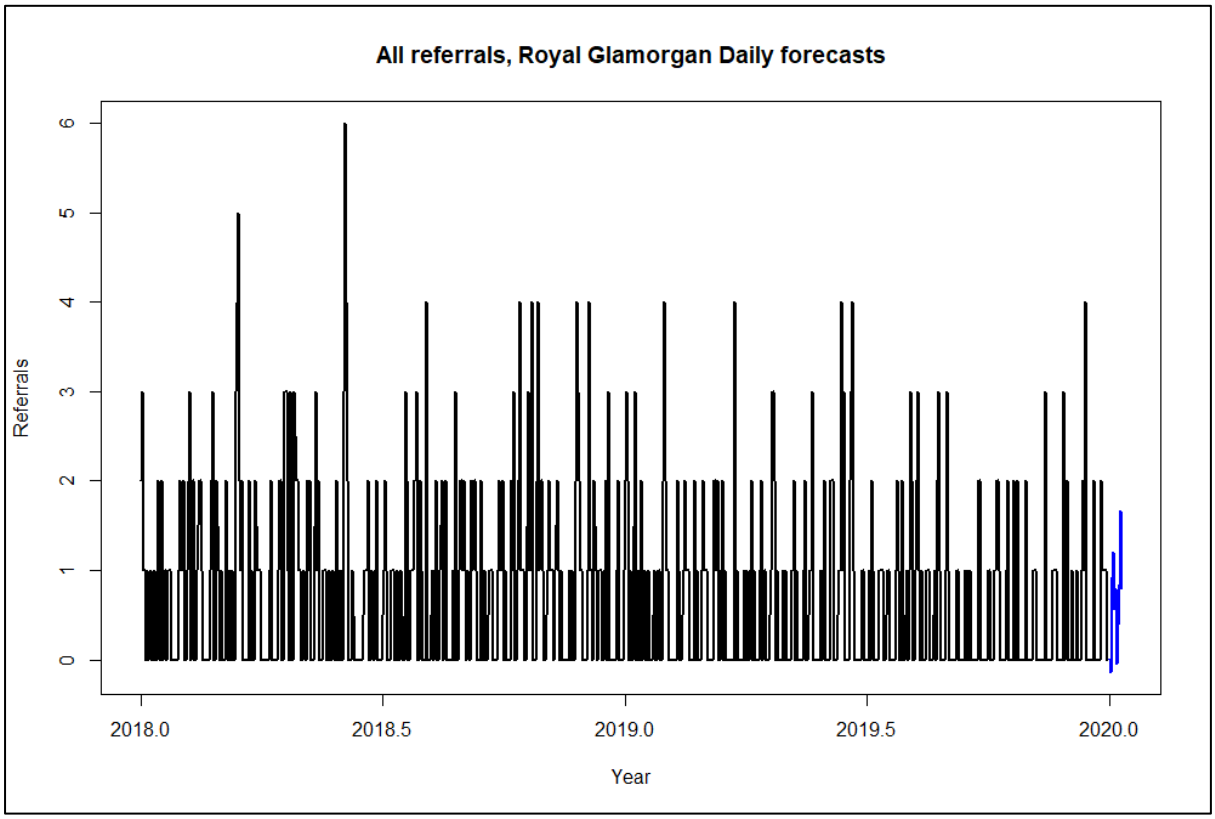

Figure 3.4: forecast- daily referrals (RGH).

3.3 Results: all referrals, Prince Charles weekly

Analysis- Prince Charles referrals, weekly

Table 3.9: forecasts from each method- referrals (PCH), weekly.

Time series method	Forecast (2019)									
Referrals- Prince Charles, weekly	w/c 04/06	w/c 11/06	w/c 18/06	w/c 25/06	w/c 02/07	w/c 09/07	w/c 16/07	w/c 23/07	w/c 30/07	w/c 06/08
Naïve	11	11	11	11	11	11	11	11	11	11
Seasonal Naïve	3	10	9	7	13	7	8	7	8	3
SES	6.92	6.92	6.92	6.92	6.92	6.92	6.92	6.92	6.92	6.92
Holt-linear	8.05	8.20	8.35	8.50	8.65	8.79	8.94	9.09	9.24	9.39
Holt-Winters	5.20	6.28	8.82	6.75	7.74	7.25	7.80	3.61	4.35	7.87
Simple linear regression	6.46	6.45	6.44	6.43	6.42	6.40	6.39	6.38	6.37	6.36
ARIMA	6.92	6.92	6.92	6.92	6.92	6.92	6.92	6.92	6.92	6.92
SSA	5.81	5.80	5.78	5.73	5.74	5.68	5.70	5.63	5.65	5.59
ANN	5.64	7.66	8.05	5.55	7.06	4.78	7.01	4.64	6.43	6.91

Table 3.10: error statistics- weekly referrals (PCH) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Prince Charles, weekly training set					
Naïve	2.58	11.47	3.39	-18.02	48.83
Seasonal Naïve	2.76	11.08	3.33	-38.80	61.89
SES	1.99	6.21	2.49	-23.81	44.24
Holt-linear	2.03	6.57	2.56	-21.91	43.03
Holt-Winters	1.82	5.30	2.30	-17.74	36.90
Simple linear regression	1.97	6.14	2.48	-23.61	43.92
ARIMA	1.99	6.21	2.49	-23.80	44.24
SSA	1.89	5.90	2.43	-21.09	41.25
ANN	1.70	4.11	2.03	-23.23	41.75

Table 3.11: error statistics- weekly referrals (PCH) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Prince Charles, weekly test set					
Naïve	3.90	17.70	4.21	35.45	35.45
Seasonal Naïve	2.60	11.60	3.41	-20.36	52.97
SES	1.11	2.52	1.59	-2.60	16.13
Holt-linear	2.10	5.31	2.31	18.35	24.04
Holt-Winters	1.81	3.92	1.98	-14.11	30.85
Simple linear regression	1.21	2.96	1.72	-10.75	18.81
ARIMA	1.12	2.52	1.59	-2.60	16.13
SSA	1.54	4.42	2.10	-24.35	26.90
ANN	0.95	2.10	1.45	-14.22	14.66

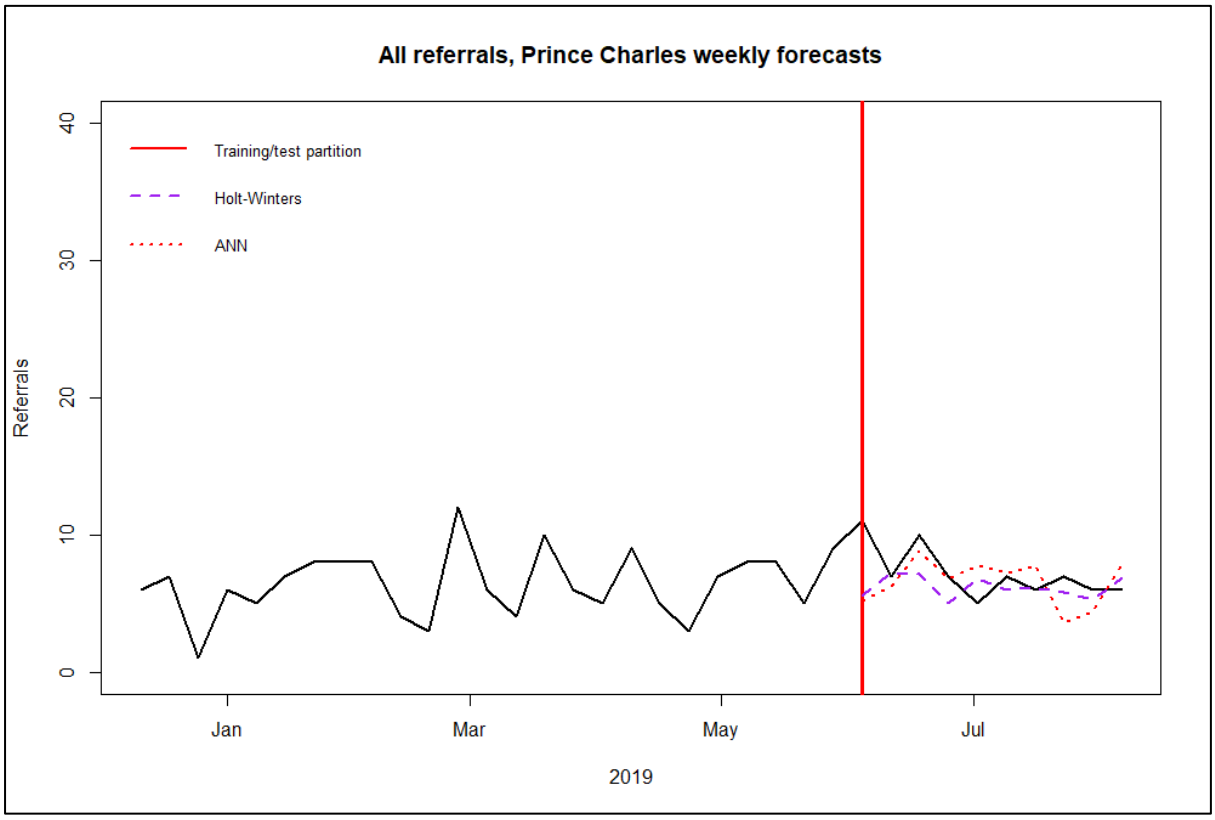

Figure 3.5: forecast- weekly referrals (PCH).

We endorse ANN method.

Table 3.12: ten-week forecast- weekly referrals (PCH).

Week commencing	Forecast (ANN)
03/12/2019	5.72
10/12/2019	6.99
17/12/2019	6.18
24/12/2019	6.36
31/12/2019	5.73
07/01/2020	6.09
14/01/2020	5.94
21/01/2020	5.98
28/01/2020	6.33
04/02/2020	6.99

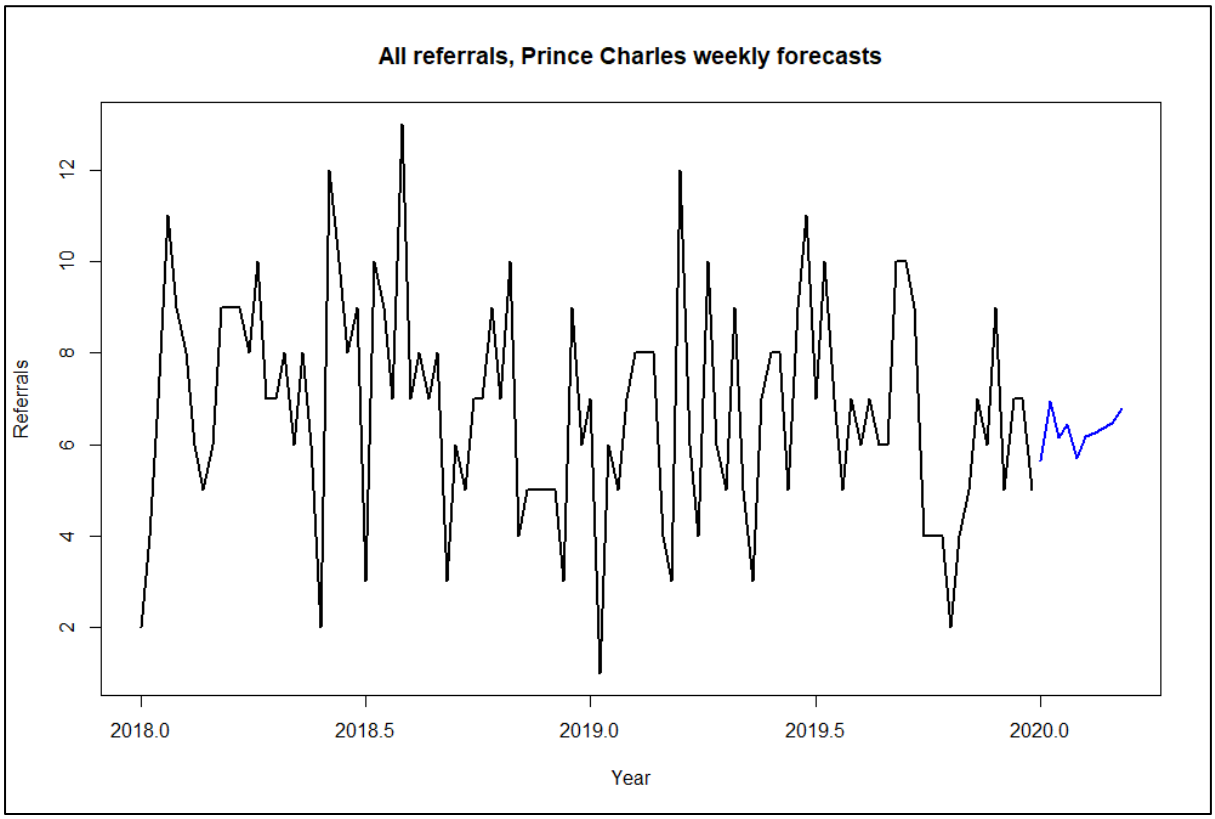

Figure 3.6: forecast- weekly referrals (PCH).

3.4 Results: all referrals, Royal Glamorgan weekly

Analysis- Royal Glamorgan referrals, weekly

Table 3.13: forecasts from each method- referrals (RGH), weekly.

Time series method	Forecast (2019)									
Referrals- Royal Glamorgan, weekly	w/c 04/06	w/c 11/06	w/c 18/06	w/c 25/06	w/c 02/07	w/c 09/07	w/c 16/07	w/c 23/07	w/c 30/07	w/c 06/08
Naïve	2	2	2	2	2	2	2	2	2	2
Seasonal Naïve	5	3	6	10	4	5	5	6	7	8
SES	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57
Holt-linear	5.37	5.37	5.37	5.37	5.36	5.36	5.36	5.35	5.35	5.35
Holt-Winters	4.72	5.84	4.82	4.40	4.32	5.31	5.81	5.41	6.42	2.37
Simple linear regression	5.13	5.12	5.11	5.10	5.09	5.07	5.06	5.05	5.04	5.03
ARIMA	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57
SSA	5.14	5.26	5.14	5.20	5.14	5.15	5.13	5.10	5.11	5.07
ANN	5.34	5.66	4.35	10.08	4.57	4.17	4.37	4.57	5.12	5.89

Table 3.14: error statistics- weekly referrals (RGH) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Royal Glamorgan, weekly training set					
Naïve	2.95	14.68	3.83	-Inf	Inf
Seasonal Naïve	2.76	14.52	3.81	-65.12	91.43
SES	2.17	6.89	2.62	-Inf	Inf
Holt-linear	2.15	6.94	2.63	-Inf	Inf
Holt-Winters	1.93	6.59	2.57	-Inf	Inf
Simple linear regression	2.14	6.82	2.61	-Inf	Inf
ARIMA	2.17	6.88	2.62	-Inf	Inf
SSA	2.10	6.36	2.52	-Inf	Inf
ANN	1.29	2.70	1.64	-16.32	31.15

Table 3.15: error statistics- weekly referrals (RGH) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Royal Glamorgan, weekly test set					
Naïve	2.70	10.10	3.18	-135	135
Seasonal Naïve	2.80	10.60	3.26	7.20	47.87
SES	1.73	3.57	1.89	15.67	31.03
Holt-linear	1.65	3.25	1.80	12.35	30.71
Holt-Winters	1.56	3.54	1.88	0.82	31.67
Simple linear regression	1.54	2.95	1.72	7.48	30.31
ARIMA	1.73	3.57	1.89	15.67	31.03
SSA	1.57	3	1.73	8.68	30.54
ANN	1.73	3.85	1.96	-1.48	37.35

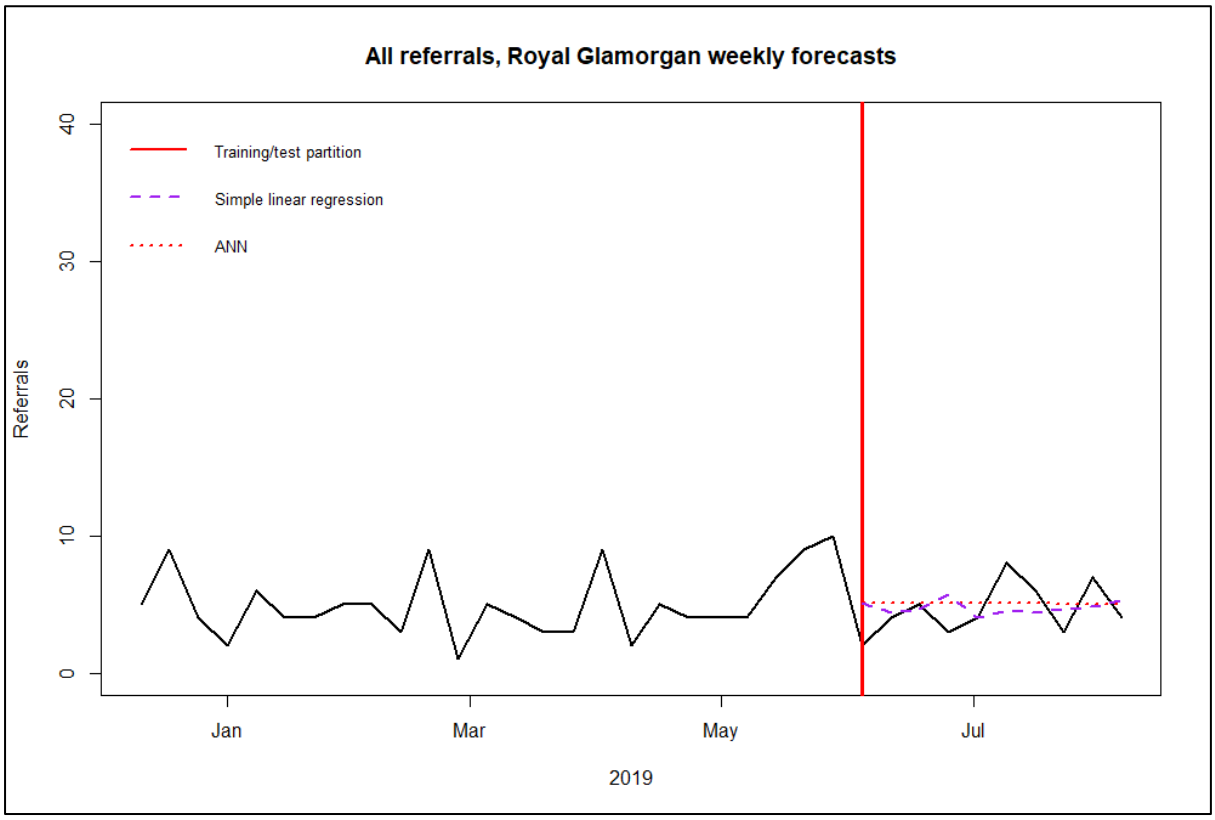

Figure 3.7: forecast- weekly referrals (RGH).

We endorse SLR method.

Table 3.16: ten-week forecast- weekly referrals (RGH).

Week commencing	Forecast (SLR)
03/12/2019	4.18
10/12/2019	4.16
17/12/2019	4.14
24/12/2019	4.12
31/12/2019	4.10
07/01/2020	4.08
14/01/2020	4.06
21/01/2020	4.04
28/01/2020	4.01
04/02/2020	3.99

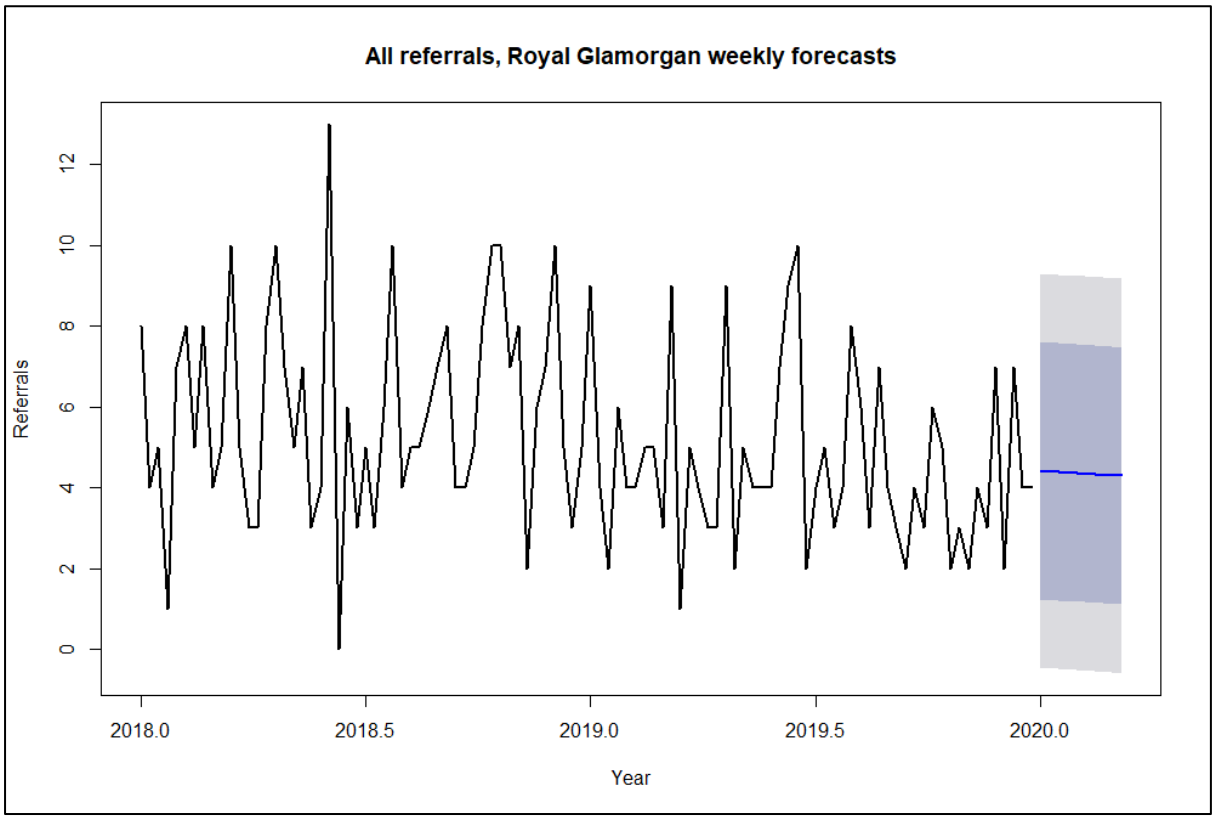

Figure 3.8: forecast- weekly referrals (RGH).

3.5 Results: all referrals, Prince Charles monthly

Analysis- Prince Charles referrals, monthly

Table 3.17: forecasts from each method- referrals (PCH), monthly.

Time series method	Forecast		
Referrals- Prince Charles, monthly	July 2019	Aug 2019	Sept 2019
Naïve	31	31	31
Seasonal Naïve	32	43	29
SES	29.70	29.70	29.70
Holt-linear	26.52	26.16	25.79
Holt-Winters	19.32	26.81	25.38
Simple linear regression	26.42	26.05	25.69
ARIMA	28.45	28.45	28.45
SSA	23.48	22.95	22.99
ANN	23.56	31.56	25.52

Table 3.18: error statistics- monthly referrals (PCH) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Prince Charles, monthly training set					
Naïve	6.31	59.69	7.73	-2.94	22.15
Seasonal Naïve	4.8	34.80	5.90	-11.34	17.40
SES	4.28	29.74	5.45	-3.49	15.11
Holt-linear	4.08	26.85	5.18	-3.42	14.40
Holt-Winters	2.79	17.24	4.15	-1.49	10.05
Simple linear regression	4.05	26.53	5.15	-3.11	14.23
ARIMA	1.99	6.21	2.49	-23.80	44.24
SSA	3.68	21.20	4.60	-0.80	12.61
ANN	0	0	0	0	0.01

Table 3.19: error statistics- monthly referrals (PCH) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Prince Charles, monthly test set					
Naïve	6.20	61.80	7.86	4.52	20
Seasonal Naïve	8.20	81.40	9.02	7.95	25.11
SES	6.06	59.85	7.74	0.35	20.40
Holt-linear	7.58	70.52	8.40	-14.49	29.33
Holt-Winters	10.47	178.63	13.37	-18.13	40.69
Simple linear regression	7.64	71.33	8.44	-14.98	29.70
ARIMA	6.31	61.17	7.82	-4.06	22.19
SSA	9.44	102.64	10.13	-29.15	41.11
ANN	9.70	115.32	10.74	-14.63	36.98

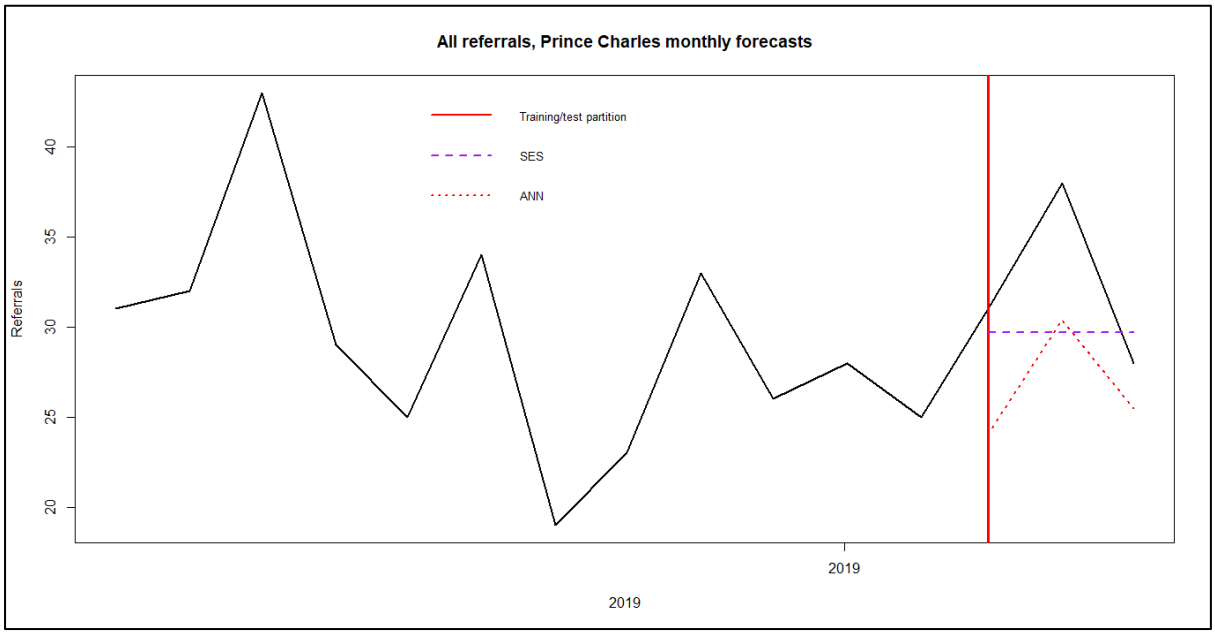

Figure 3.9: forecast- monthly referrals (PCH).

We endorse SES method.

Table 3.20: three-month forecast- monthly referrals (PCH).

Month	Forecast (SES)
Dec 2019	29.48
Jan 2020	29.48
Feb 2020	29.48

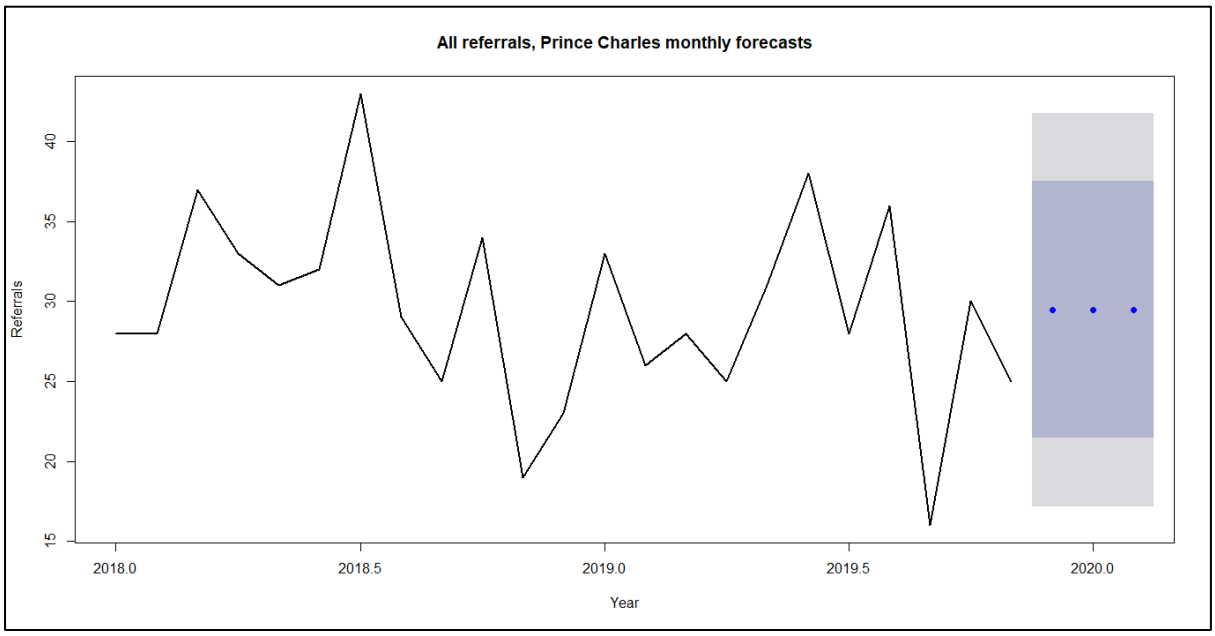

Figure 3.10: forecast- monthly referrals (PCH).

3.6 Results: all referrals, Royal Glamorgan monthly

Analysis- Royal Glamorgan referrals, monthly

Table 3.21: forecasts from each method- referrals (RGH), monthly.

Time series method	Forecast		
Referrals- Royal Glamorgan, monthly	July 2019	Aug 2019	Sept 2019
Naïve	30	30	30
Seasonal Naïve	14	25	30
SES	24.47	24.47	24.47
Holt-linear	22.95	22.78	22.60
Holt-Winters	17.19	16.15	17.59
Simple linear regression	23.05	22.89	22.73
ARIMA	24.42	24.47	24.47
SSA	23.19	22.93	20.35
ANN	24.31	21.42	25.83

Table 3.22: error statistics- monthly referrals (RGH) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Royal Glamorgan, monthly training set					
Naïve	6.63	67.25	8.20	-4.54	29.19
Seasonal Naïve	4.80	36.40	6.03	-27.17	27.17
SES	4.44	32.49	5.70	-6.25	20.41
Holt-linear	4.30	31.91	5.65	-6.30	19.74
Holt-Winters	3.23	18.78	4.33	-2.67	15.92
Simple linear regression	4.31	31.89	5.65	-6.10	19.73
ARIMA	4.45	32.48	5.70	-6.25	20.42
SSA	3.43	20.08	4.48	-6.80	16.25
ANN	0.24	0.14	0.38	-0.23	1.38

Table 3.23: error statistics- monthly referrals (RGH) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Referrals- Royal Glamorgan, monthly test set					
Naïve	12.20	172.20	13.12	40.67	40.67
Seasonal Naïve	8.40	129.20	11.37	24.10	27.30
SES	7.68	67.86	8.24	27.26	31.39
Holt-linear	6.49	46.26	6.80	21.25	28.67
Holt-Winters	3.72	28.30	5.32	-4.75	22.39
Simple linear regression	6.58	47.56	6.90	21.71	28.89
ARIMA	7.67	67.62	8.22	27.23	31.37
SSA	5.13	32.21	5.68	16.54	23.64
ANN	11.50	182.70	13.52	32.93	38.24

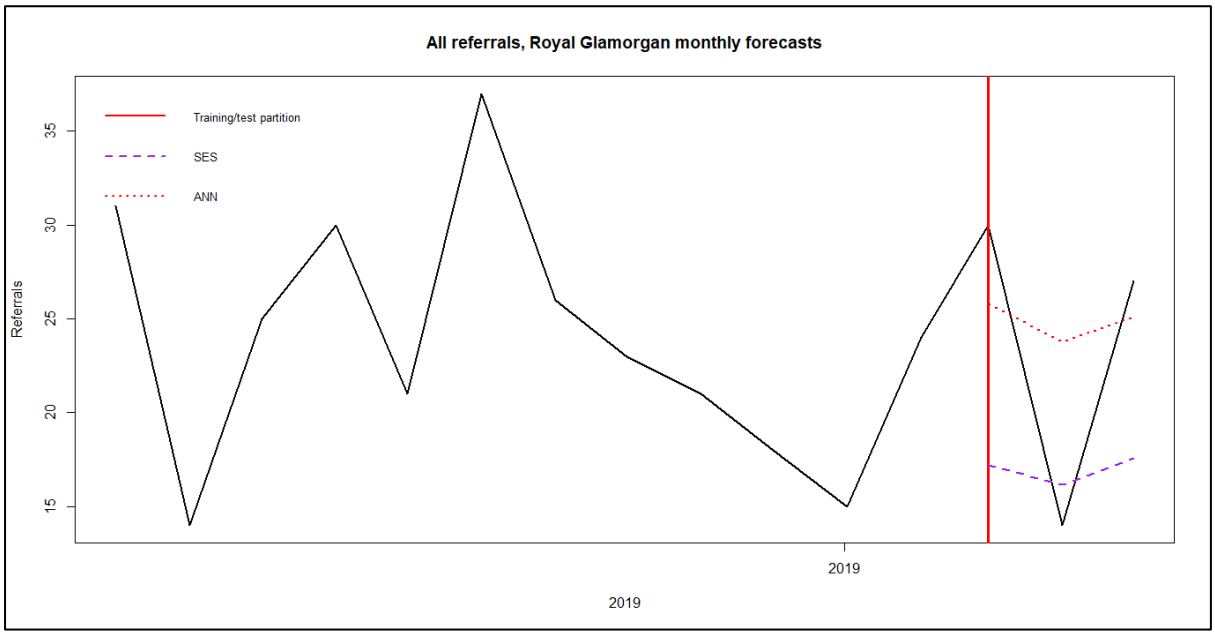

Figure 3.11: forecast- monthly referrals (RGH).

We endorse Holt-Winters method.

Table 3.24: three-month forecast- monthly referrals (RGH).

Month	Forecast (Holt-Winters)
Dec 2019	16.94
Jan 2020	14.94
Feb 2020	11.94

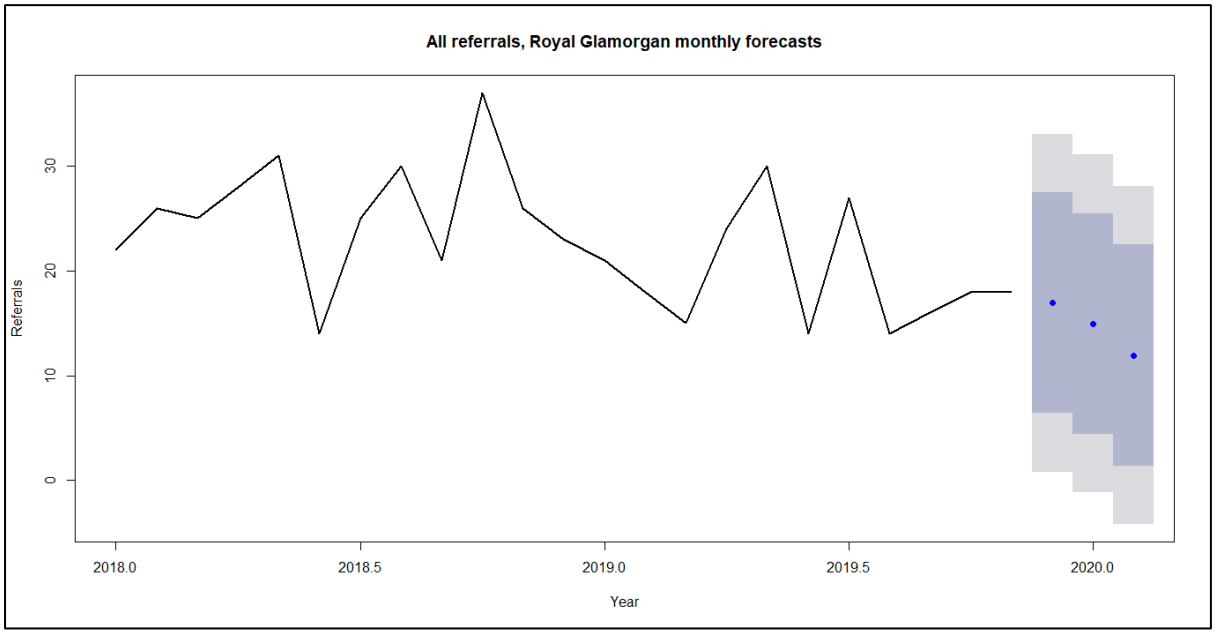

Figure 3.12: forecast- monthly referrals (RGH).

3.7 Results: CT tests weekly

Analysis- CT tests, weekly.

Table 3.25: forecasts from each method- tests (CT), weekly.

Time series method	Forecast (2019)									
Tests- CT tests, weekly	w/c 06/06	w/c 13/06	w/c 20/06	w/c 27/06	w/c 04/07	w/c 11/07	w/c 18/07	w/c 25/07	w/c 01/08	w/c 08/08
Naïve	9	9	9	9	9	9	9	9	9	9
Seasonal Naïve	13	8	8	15	12	12	16	12	9	11
SES	7.52	7.52	7.52	7.52	7.52	7.52	7.52	7.52	7.52	7.52
Holt-linear	7.76	7.85	7.94	8.04	8.13	8.22	8.31	8.40	8.50	8.59
Holt-Winters	8.32	6.96	13.60	10.37	8.26	12.19	11.77	9.14	8.25	10.39
Simple linear regression	8.84	8.86	8.87	8.88	8.89	8.91	8.92	8.93	8.95	8.96
ARIMA	7.53	7.53	7.53	7.53	7.53	7.53	7.53	7.53	7.53	7.53
SSA	6.95	6.85	6.77	6.70	6.64	6.60	6.58	6.56	6.55	6.55
ANN	8.54	5.54	6.14	8.45	8.21	5.56	7.24	7.81	6.84	6.10

Table 3.26: error statistics- weekly tests (CT) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- CT tests, weekly training set					
Naïve	2.92	13.88	3.73	-Inf	Inf
Seasonal Naïve	4.91	31.53	5.62	-19.74	76.02
SES	2.44	10.25	3.20	-Inf	Inf
Holt-linear	2.44	10.22	3.20	-Inf	Inf
Holt-Winters	2.04	6.87	2.62	Inf	Inf
Simple linear regression	3.44	18.04	4.25	-Inf	Inf
ARIMA	2.44	10.25	3.20	-Inf	Inf
SSA	3.07	13.89	3.73	-Inf	Inf
ANN	1.14	1.89	1.37	-6.77	18.92

Table 3.27: error statistics- weekly tests (CT) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- CT tests, weekly test set					
Naïve	4.40	24.60	4.96	48.89	48.49
Seasonal Naïve	7	58.68	7.66	59.02	59.02
SES	3.21	13.76	3.71	38.82	42.76
Holt-linear	3.82	18.58	4.31	43.28	46.47
Holt-Winters	5.46	38.69	6.22	51.31	52.95
Simple linear regression	4.33	23.81	4.88	48.28	48.63
ARIMA	3.23	13.84	3.72	38.93	42.82
SSA	2.49	9.24	3.04	31.44	37.33
ANN	2.37	11.97	3.46	30.27	42.22

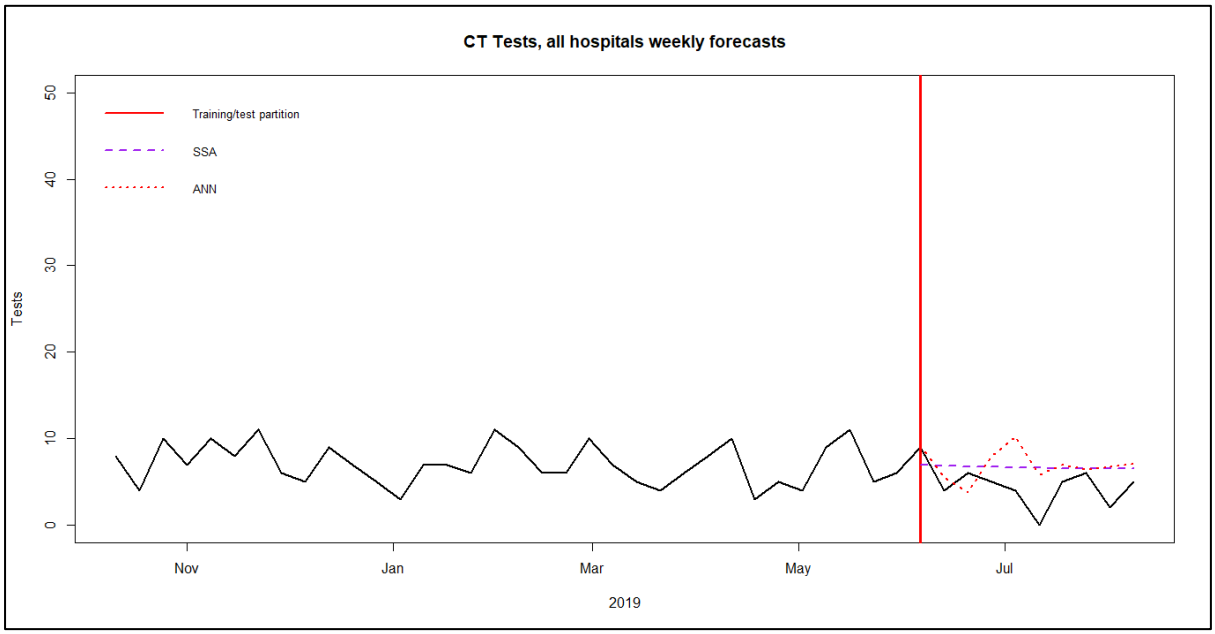

Figure 3.13: forecast- weekly tests (CT).

We endorse SSA method.

Table 3.28: ten-week forecast- weekly tests (CT).

Week commencing	Forecast (SSA)
26/12/2019	11.86
02/01/2020	10.10
09/01/2020	12.04
16/01/2020	11.92
23/01/2020	9.75
30/01/2020	11.28
06/02/2020	11.93
13/02/2020	9.50
20/02/2020	10.56
27/02/2020	11.78

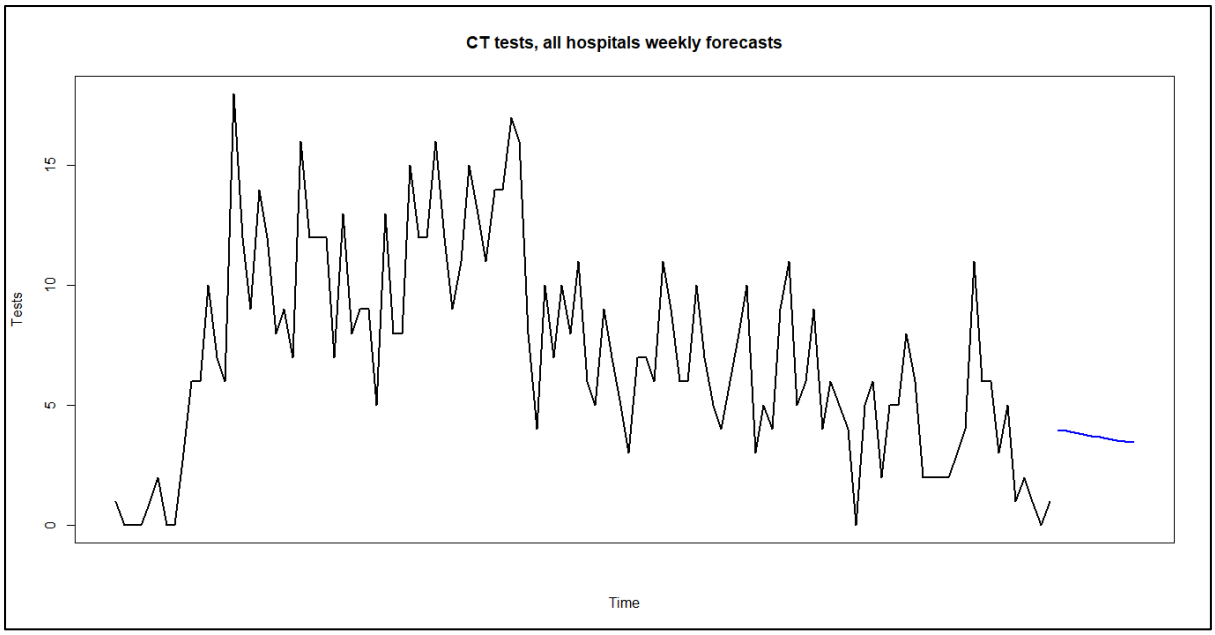

Figure 3.14: forecast- weekly tests (CT).

3.8 Results: PET tests weekly

Analysis- PET tests, weekly

Table 3.29: forecasts from each method- tests (PET), weekly.

Time series method	Forecast (2019)									
Tests- PET tests, weekly	11/06	12/06	13/06	14/06	15/06	16/06	17/06	18/06	19/06	20/06
Naïve	1	1	1	1	1	1	1	1	1	1
Seasonal Naïve	4	2	2	3	1	2	3	3	5	2
SES	1.95	1.95	1.95	1.95	1.95	1.95	1.95	1.95	1.95	1.95
Holt-linear	1.79	1.73	1.67	1.61	1.56	1.50	1.44	1.38	1.32	1.27
Holt-Winters	1.73	2.48	1.97	2.66	3.24	0.90	2.15	3.57	1.26	3.60
Simple linear regression	3.21	3.23	3.24	3.25	3.26	3.28	3.29	3.30	3.31	3.33
ARIMA	1.97	1.69	1.77	1.75	1.76	1.76	1.76	1.76	1.76	1.76
SSA	2.64	2.94	2.40	2.96	2.52	2.66	2.83	2.37	2.92	2.42
ANN	3.46	5.81	2.83	0	3.50	2.76	2.60	2.84	1.25	2.34

Table 3.30: error statistics- weekly tests (PET) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- PET tests, weekly training set					
Naïve	1.94	6.71	2.59	-Inf	Inf
Seasonal Naïve	2.66	10.53	3.25	-Inf	Inf
SES	1.38	3.37	1.84	-Inf	Inf
Holt-linear	1.43	3.39	1.84	Inf	Inf
Holt-Winters	1.35	2.84	1.68	Inf	Inf
Simple linear regression	1.55	3.70	1.92	-Inf	Inf
ARIMA	1.32	3.17	1.78	-Inf	Inf
SSA	1.46	3.07	1.75	-Inf	Inf
ANN	0	0	0.01	Inf	Inf

Table 3.31: error statistics- weekly tests (PET) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- PET tests, weekly test set					
Naïve	4.40	24.60	4.96	48.89	48.49
Seasonal Naïve	7	58.68	7.66	59.02	59.02
SES	3.21	13.76	3.71	38.82	42.76
Holt-linear	3.82	18.58	4.31	43.28	46.47
Holt-Winters	5.46	38.69	6.22	51.31	52.95
Simple linear regression	4.33	23.81	4.88	48.28	48.63
ARIMA	3.23	13.84	3.72	38.93	42.82
SSA	2.49	9.24	3.04	31.44	37.33
ANN	2.37	11.97	3.46	30.27	42.22

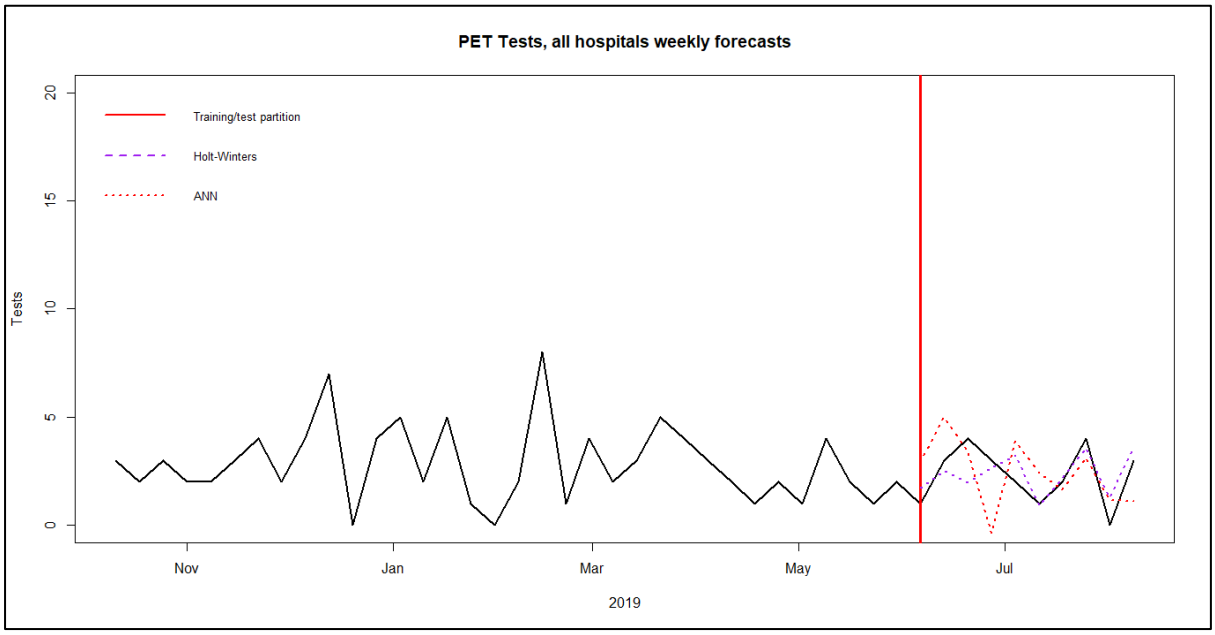

Figure 3.15: forecast- weekly tests (PET).

We endorse Holt-Winters method.

Table 3.32: ten-week forecast- weekly tests (PET).

Week commencing	Forecast (Holt-Winters)
26/12/2019	1.19
02/01/2020	0
09/01/2020	1.02
16/01/2020	1.17
23/01/2020	0
30/01/2020	2.82
06/02/2020	0
13/02/2020	1.73
20/02/2020	1.33
27/02/2020	0.83

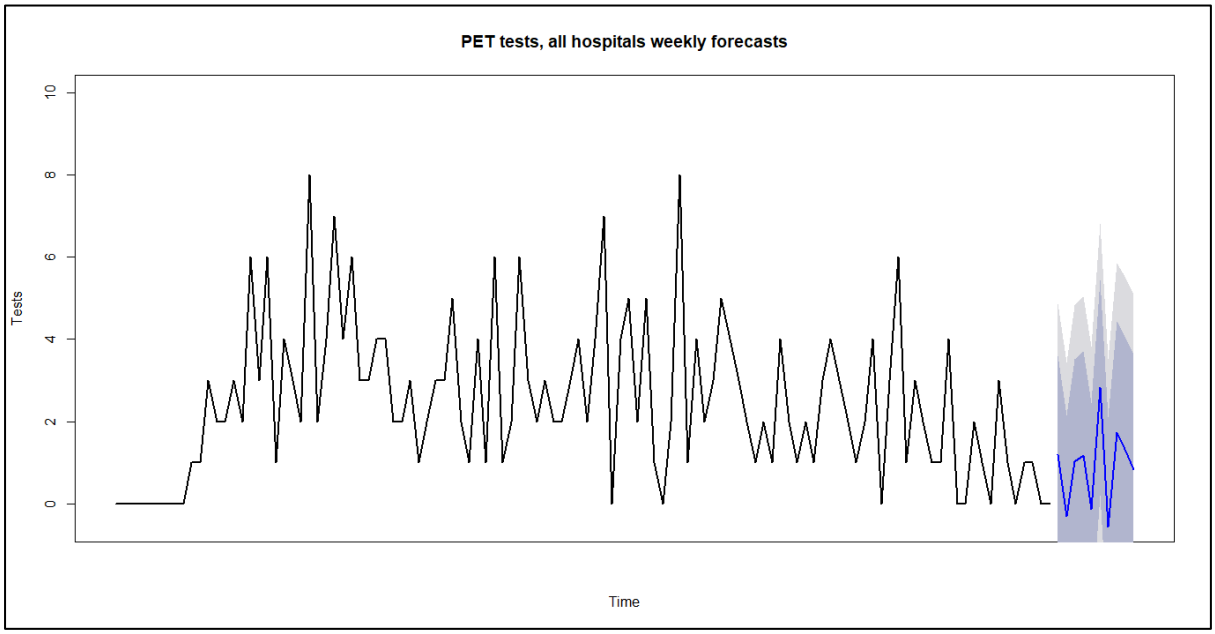

Figure 3.16: forecast- weekly tests (PET).

3.9 Results: Other tests weekly

Analysis- Other tests, weekly

Table 3.33: forecasts from each method- tests (Other) weekly.

Time series method	Forecast (2019)									
Tests- Other tests, weekly	11/06	12/06	13/06	14/06	15/06	16/06	17/06	18/06	19/06	20/06
Naïve	6	6	6	6	6	6	6	6	6	6
Seasonal Naïve	6	4	2	3	6	9	4	6	5	8
SES	4.94	4.94	4.94	4.94	4.94	4.94	4.94	4.94	4.94	4.94
Holt-linear	5.13	5.19	5.25	5.31	5.37	5.43	5.49	5.55	5.61	5.67
Holt-Winters	6.14	6.25	5.87	6.47	5.86	5.28	7.58	10.69	7.74	5.33
Simple linear regression	5.75	5.78	5.81	5.84	5.87	5.89	5.92	5.95	5.98	6.01
ARIMA	4.96	4.96	4.96	4.96	4.96	4.96	4.96	4.96	4.96	4.96
SSA	4.70	4.66	4.65	4.66	4.66	4.64	4.60	4.56	4.54	4.54
ANN	3.82	2.79	0	9	9.14	4.13	4.40	5.68	5.35	5.22

Table 3.34: error statistics- weekly tests (Other) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- Other tests, weekly training set					
Naïve	2.14	8.29	2.88	-18.91	57.19
Seasonal Naïve	3.06	14	3.74	-4.38	83.44
SES	1.89	5.80	2.41	-Inf	Inf
Holt-linear	1.90	5.76	2.40	-Inf	Inf
Holt-Winters	1.64	4.27	2.07	Inf	Inf
Simple linear regression	2.12	7.40	2.72	-Inf	Inf
ARIMA	1.88	5.80	2.41	-20.43	54.38
SSA	2.02	6.22	2.49	-Inf	Inf
ANN	1.05	1.50	1.23	-17.08	33.29

Table 3.35: error statistics- weekly tests (Other) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- Other tests, weekly test set					
Naïve	1.90	5.90	2.43	21.67	31.67
Seasonal Naïve	2.20	9.42	3.07	-7.86	47.97
SES	1.51	4.28	2.07	4.83	30.62
Holt-linear	1.70	4.71	2.17	12.93	31.16
Holt-Winters	2.76	11.02	3.32	26.68	40.47
Simple linear regression	1.87	5.62	2.37	20.07	31.65
ARIMA	1.51	4.28	2.07	5.22	30.41
SSA	1.57	4.28	2.07	-1.85	34.01
ANN	5.43	48.44	6.96	50.67	101.55

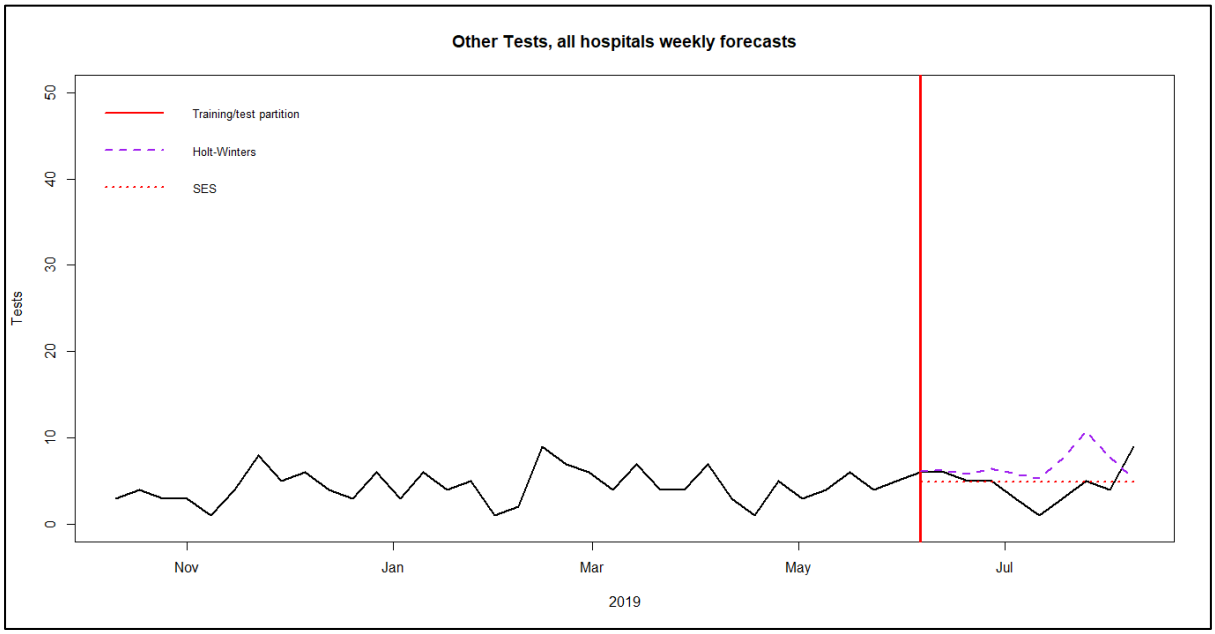

Figure 3.17: forecast- weekly tests (Other).

We endorse SES method.

Table 3.36: ten-week forecast- weekly tests (Other).

Week commencing	Forecast (SES)
26/12/2019	1.69
02/01/2020	1.69
09/01/2020	1.69
16/01/2020	1.69
23/01/2020	1.69
30/01/2020	1.69
06/02/2020	1.69
13/02/2020	1.69
20/02/2020	1.69
27/02/2020	1.69

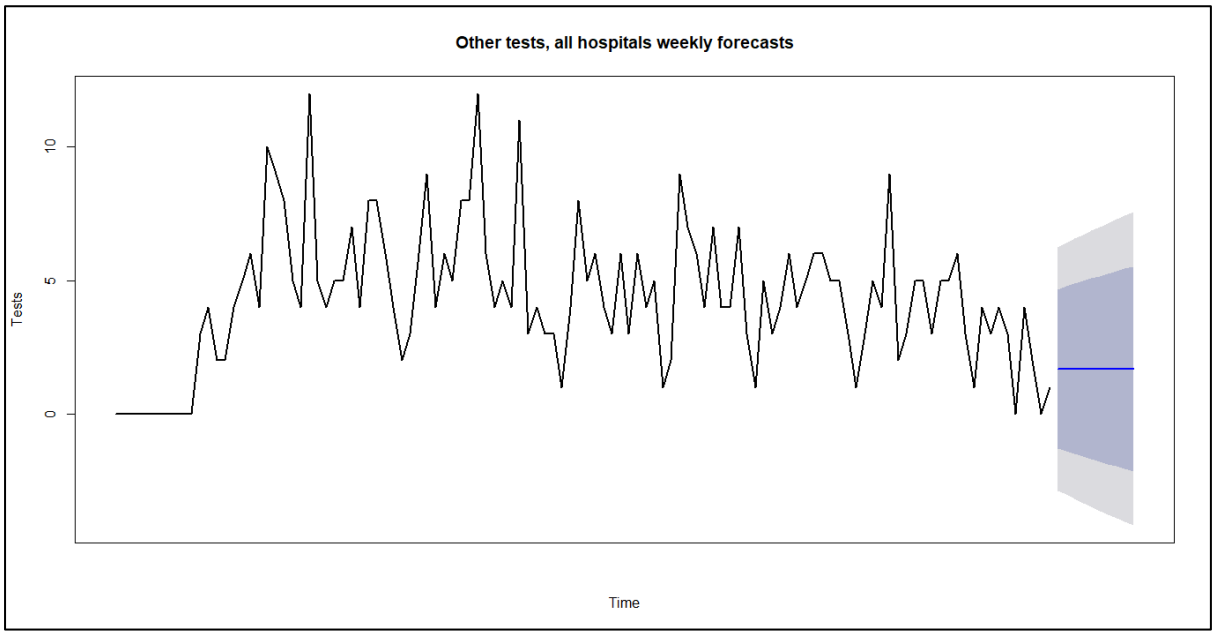

Figure 3.18: forecast- weekly tests (Other).

3.10 Results: CT tests monthly

Analysis- CT tests, monthly

Table 3.37: forecasts from each method- tests (CT) monthly.

Time series method	Forecast		
Tests- CT tests, monthly	May 2019	June 2019	July 2019
Naïve	32	32	32
Seasonal Naïve	36	60	52
SES	32	32	32
Holt-linear	33.67	35.35	37.03
Holt-Winters	32.41	32.65	32.90
Simple linear regression	39.27	39.59	39.92
ARIMA	32.26	35.44	35.44
SSA	31.52	30.34	29.43
ANN	29.30	30.02	30.13

Table 3.38: error statistics- monthly tests (CT) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- CT tests, monthly training set					
Naïve	8.56	132.89	11.53	3.16	21.22
Seasonal Naïve	21.60	466.71	21.60	-8.10	61.77
SES	8.11	125.89	11.22	3.01	20.11
Holt-linear	8.89	130.17	11.41	-35.94	55.43
Holt-Winters	9.19	162.20	12.74	2.42	22.59
Simple linear regression	12.45	234.31	15.31	-162.94	187.11
ARIMA	8.57	128.21	11.32	-107.93	126.26
SSA	9	155.22	12.46	-169.66	178.47
ANN	0.61	1.17	1.08	-0.14	2.07

Table 3.39: error statistics- monthly tests (CT) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- CT tests, monthly test set					
Naïve	10.86	155.14	12.46	33.93	33.93
Seasonal Naïve	24.14	737.57	27.16	50.50	50.50
SES	10.86	155.14	12.46	33.93	33.93
Holt-linear	17.57	383.45	19.58	44.03	44.03
Holt-Winters	10.86	155.14	12.46	33.93	33.93
Simple linear regression	19.09	407.28	20.18	47.28	47.28
ARIMA	14.13	240.84	15.52	39.91	39.91
SSA	14.41	243.94	15.62	-213.13	213.13
ANN	9.35	109.31	10.46	27.68	31.12

**Holt-Winters could not estimate seasonality due to insufficient data.*

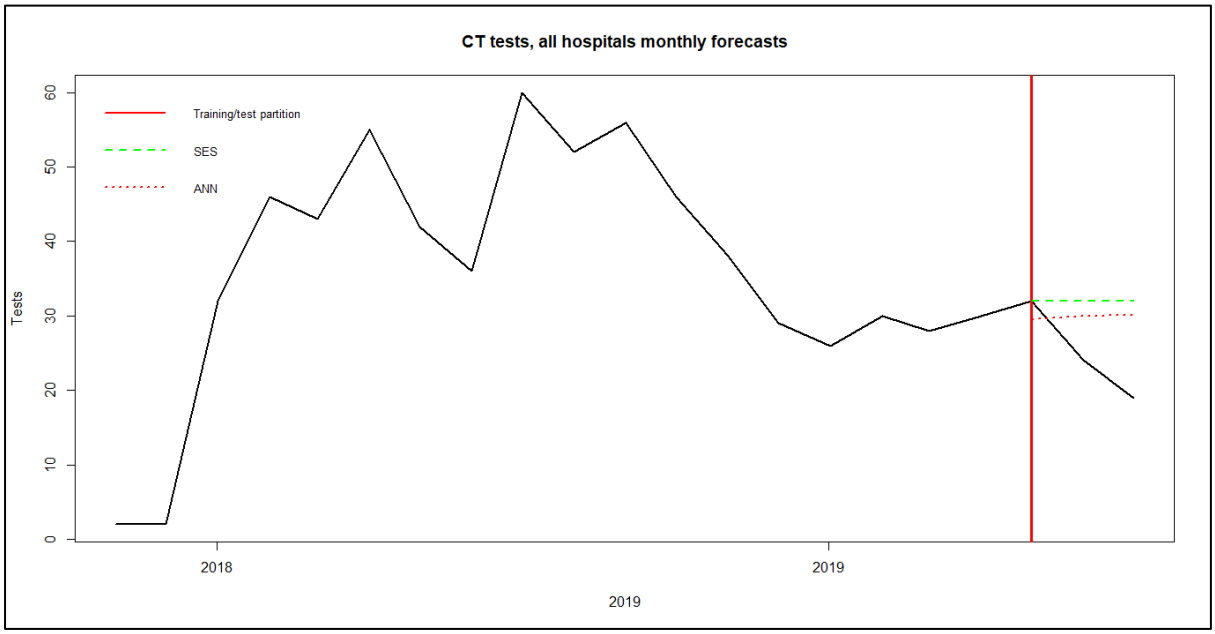

Figure 3.19: forecast- monthly tests (CT).

We endorse ANN method

Table 3.40: three-month forecast- monthly tests (CT).

Month	Forecast (ANN)
Jan 20	0
Feb 20	0
Mar 20	0

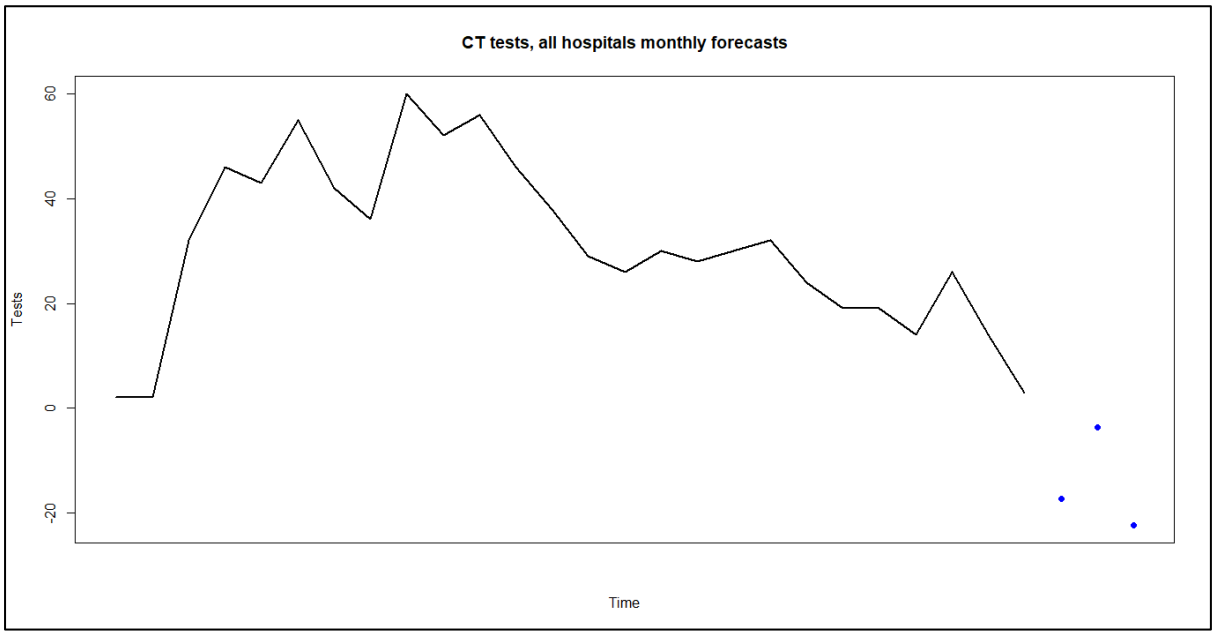

Figure 3.20: forecast- monthly tests (CT).

3.11 Results: PET tests monthly

Analysis- PET tests, monthly

Table 3.41: forecasts from each method- tests (PET) monthly.

Time series method	Forecast		
Tests- PET tests, monthly	May 2019	June 2019	July 2019
Naïve	8	8	8
Seasonal Naïve	12	11	11
SES	8.69	8.69	8.69
Holt-linear	9.50	10.10	10.70
Holt-Winters	6.27	3.10	0
Simple linear regression	14.75	15.06	15.37
ARIMA	10.52	11.45	11.45
SSA	12.47	11.97	11.88
ANN	17.96	10.92	17.24

Table 3.42: error statistics- monthly tests (PET) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- PET tests, monthly training set					
Naïve	3.78	23.33	4.83	1.15	33.37
Seasonal Naïve	8.29	96.57	9.83	0.26	75.59
SES	3.71	21.36	4.62	-Inf	Inf
Holt-linear	3.64	21.07	4.59	Inf	-Inf
Holt-Winters	5.20	40.17	6.34	-2.53	43.84
Simple linear regression	3.73	25.71	5.07	-Inf	Inf
ARIMA	3.70	20.69	4.55	-Inf	Inf
SSA	3.18	19.31	4.39	-Inf	Inf
ANN	0.12	0.02	0.13	-0.08	0.98

Table 3.43: error statistics- monthly tests (PET) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- PET tests, monthly test set					
Naïve	2.29	7.14	2.67	-3.57	28.57
Seasonal Naïve	4.29	25.71	5.07	31.46	34.05
SES	2.19	7.22	2.69	4.63	25.18
Holt-linear	3.38	21.80	4.67	24.26	27.74
Holt-Winters	2.19	7.22	2.69	4.52	25.22
Simple linear regression	7.40	64.34	8.02	46.65	46.65
ARIMA	3.19	16.28	4.03	26.75	28.12
SSA	3.61	18.82	4.34	30.53	30.81
ANN	4.47	28.45	5.33	29.16	29.16

**Holt-Winters could not estimate seasonality due to insufficient data.*

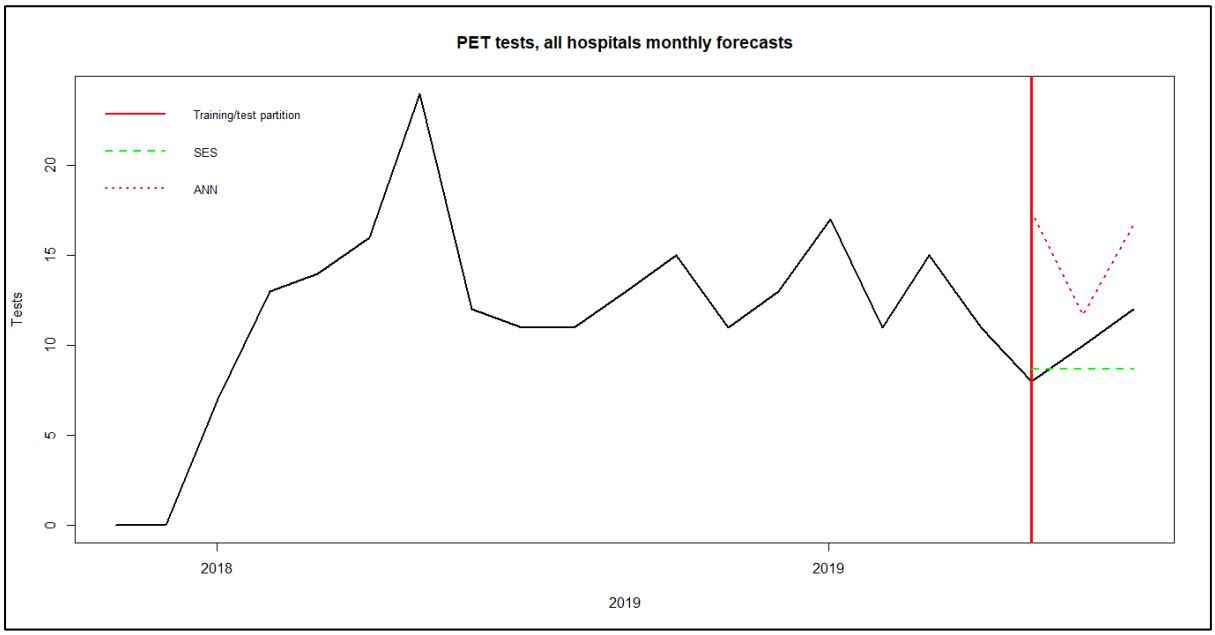

Figure 3.21: forecast- monthly tests (PET).

We endorse SES method.

Table 3.44: three-month forecast- monthly tests (PET).

Month	Forecast (SES)
Jan 20	2.36
Feb 20	2.36
Mar 20	2.36

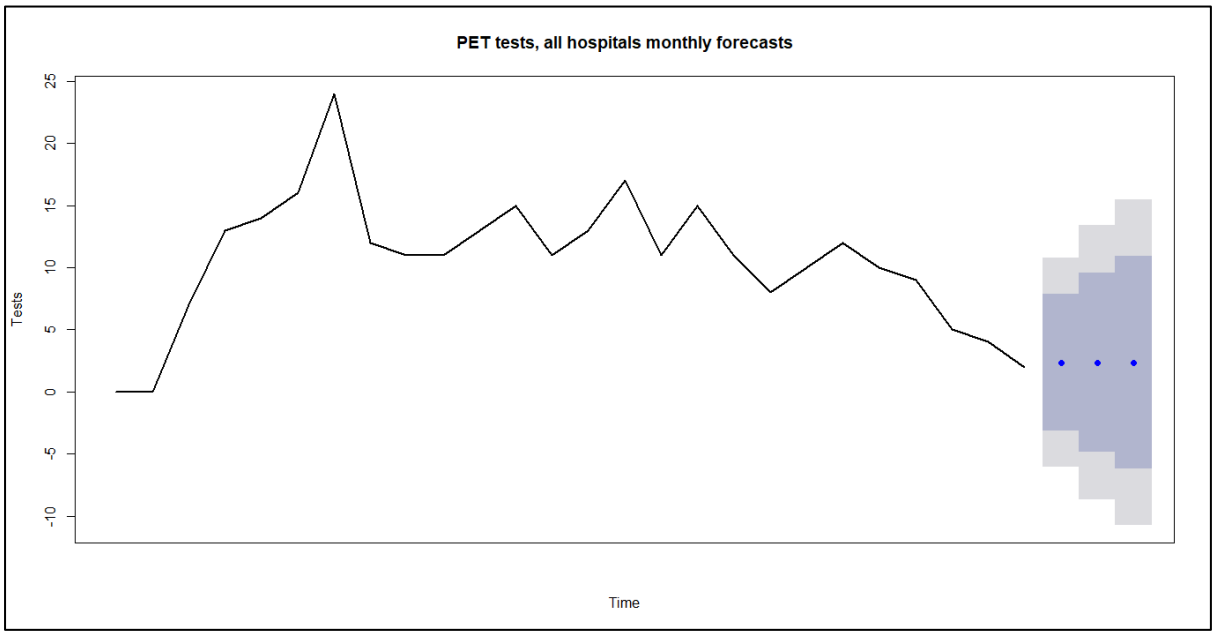

Figure 3.22: forecast- monthly tests (PET).

3.12 Results: Other tests monthly

Analysis- Other tests, monthly

Table 3.45: forecasts from each method- tests (Other) monthly.

Time series method	Forecast		
Tests- Other tests, monthly	May 2019	June 2019	July 2019
Naïve	17	17	17
Seasonal Naïve	26	24	32
SES	17.01	17.01	17.01
Holt-linear	18.26	19.29	20.33
Holt-Winters	16.98	17.11	17.25
Simple linear regression	24.93	25.46	26
ARIMA	18.01	19.20	19.20
SSA	20.11	20	19.82
ANN	19.78	19.18	19.98

Table 3.46: error statistics- monthly tests (Other) training set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- Other tests, monthly training set					
Naïve	4.72	35.61	5.97	6.07	25.88
Seasonal Naïve	10.71	147	12.12	20.96	56.08
SES	4.47	33.74	5.81	-Inf	Inf
Holt-linear	4.49	32.94	5.74	-Inf	Inf
Holt-Winters	5.13	43.20	6.57	5	26.44
Simple linear regression	6.47	63.67	7.98	-Inf	Inf
ARIMA	4.88	44.84	6.70	-Inf	Inf
SSA	4.87	51.66	7.19	-Inf	Inf
ANN	0.66	0.84	0.92	0	3.48

Table 3.47: error statistics- monthly tests (Other) test set.

Time series method	MAE	MSE	RMSE	MPE	MAPE
Tests- Other tests, monthly test set					
Naïve	3	17	4.12	-2.52	17.65
Seasonal Naïve	6.14	64.43	8.03	24.36	24.36
SES	3	16.99	4.12	-2.44	17.64
Holt-linear	4.71	45.73	6.76	16.71	20.72
Holt-Winters	14.43	229.19	15.14	-379.16	423.49
Simple linear regression	9.09	105.36	10.26	33.85	33.85
ARIMA	3.08	17.03	4.13	-4.89	18.62
SSA	3.19	17.55	4.19	9.87	16.79
ANN	2.20	6.10	2.47	-0.79	11.37

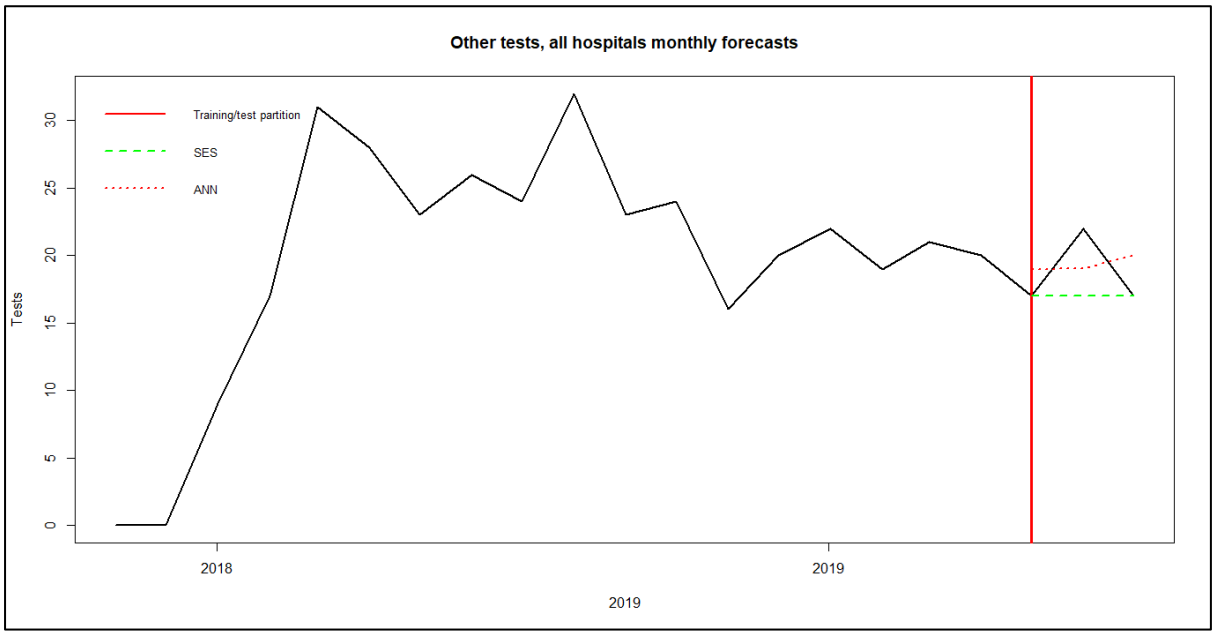

Figure 3.23: forecast- monthly tests (Other).

We endorse ANN method.

Table 3.48: three-month forecast- monthly tests (Other).

Month	Forecast (ANN)
Jan 20	7.16
Feb 20	0
Mar 20	0

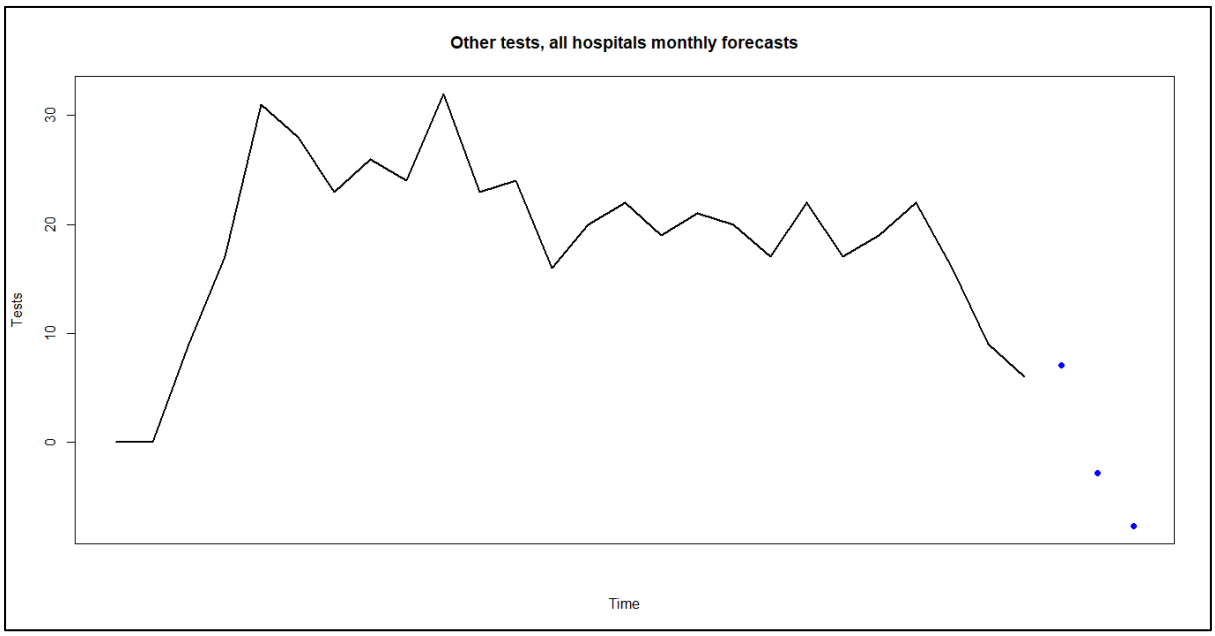

Figure 3.24: forecast- monthly tests (Other).