Force Response Deployment on Bank Holidays

Data Analytics Lab

June 2022

The aim is to deploy the appropriate numbers of Force Response officers on each Bank Holiday, according to historical and projected demand.

1 Contents

2 Introduction

Currently, Force Response deploys 138 staff each Bank Holiday shift, at a very significant cost. The staffing is the same, irrespective of whether it is Christmas morning, or August Bank Holiday, or New Year's Day.

Our aim is to determine the most efficient deployment numbers for each Bank Holiday, based on analysis of demand over the past 5.5 years (includes the pre-COVID period, to avoid misleading inference).

To accomplish this, an analysis of the numbers of incidents which are dealt with by Operational Resources (high priority incidents (P1, 2 and 3)) on bank holidays has been undertaken. This forms the bases of recommendations for the appropriate numbers of officers and the general demand profile across bank holidays.

3 Executive Summary

The number of incidents that occur on different bank holidays shift days (7am-7am) are highly variable, with the busiest bank holiday having 30-59% more incidents than the quietest. This demonstrates the need to adjust the bank holiday staff numbers. To ensure a similar level of service across the bank holidays, the staffing numbers should be in line with the expected number of P1-3 incidents.

Christmas Day is normally the quietest bank holiday in the year, followed by New Year's Day and Boxing Day (which has around 10% more incidents than Christmas). Good Friday and the Spring Bank Holidays tend to be the busiest bank holidays in the year.

In general, the proportions of incidents across the different shifts in a day, do not vary much between bank holidays, apart from Christmas and to a much lesser extent, Boxing Day, which have a higher proportion in the night shift.

The proportions of P1 (highest priority), P2 and P3 were also shown to be reasonably consistent between the bank holidays. Christmas day and Boxing Day showed higher proportions of P1 incidents, but are also two of the quietest days in the year. All bank holidays showed that nights have a higher proportion of P1 incidents (51-62%), compared to lates (44-53%), which have a higher proportion than earlies (40-45%).

A time series model was developed to assess the impact of bank holidays compared to the general trend, and also make predictions for the next 12 months of bank holidays. The fitted model showed that in general bank holidays do not have a large impact on the expected number of calls, with half the bank holidays having an effect of less than 2% when compared to the general trend. The bank holiday with the biggest negative impact was Christmas day, which gives a 12.5% reduction compared to the trend.

Using the numbers of P1-3 incidents predicted for the next 12 months of bank holidays, it is possible to adjust the staffing numbers in line with the expected number of incidents (supply matching demand). The overall numbers of staff used is the same.

These numbers can be found in the Table below:

Bank Holiday	Date	Existing Staff Number	Adjusted Staff Number	Early Adjusted Staff Number	Late Adjusted Staff Number	Night Adjusted Staff Number
Summer Bank Holiday	29/08/2022	138	146	41	70	35
Boxing Day	26/12/2022	138	133	32	62	39
Christmas Day	27/12/2022	138	115	34	55	26
New Year's Day	02/01/2023	138	117	35	56	26
Good Friday	07/04/2023	138	161	44	76	41
Easter Monday	10/04/2023	138	135	39	69	27
Early May Bank Holiday	01/05/2023	138	146	39	75	32
Spring Bank Holiday	29/05/2023	138	151	41	77	33

Notes: Shift times are; Earlies: 7am-3pm, Late: 3pm-10pm and Nights: 10pm-7am

4 Data

All P1-3 incidents have been included within the analysis, and related to the date and time that Force Response were first aware of the incident (the first time each incident is flagged as P1-3). P1-3 are those incidents which have been graded using THRIVE+ principles[1] as requiring a response within 15 minutes, 60 minutes or 24 hours respectively. Some of these incidents are dealt with by other police units which are not response (i.e. Neighborhood policing units), but probably should be dealt with by response units, or no resource is required to attend as it transpires that a call is not an incident that Force Response would deal with but the initial process has placed the incident on the Response queue.

It was a concern that Covid-19 and the related restrictions would have had an impact on Force Response demand. To take into account the different levels of restrictions, the Oxford Covid-19 Government Response Tracker[2] has been used. In general, the stringency index has been used as this is a simple index which quantifies the level of restrictions across many areas of society including workplace and school closures as well as public events and gatherings.

Bank holidays have been defined in line with government information. This means that certain bank holidays are shifted if they fall on a weekend. This includes; Christmas Day, Boxing Day and New Year's Day.

[1] Threat, Harm, Risk, Investigation, Vulnerability, Engagement, Prevention & Intervention

[2] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. (2021). "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)." Nature Human Behaviour. https://doi.org/10.1038/s41562-021-01079-8

5 Exploratory Analysis of Force Response Demand on Bank Holidays

5.1 Assessing the impact at a daily level

Initially, the impact on the total number of incidents per day shift, defined as 7am to 7am, was explored. Comparing the totals year on year in Figure 1, we can see that in general, the total number of P1-3 incidents that occurs on bank holidays (and non-bank holidays) has reduced in time. This movement seems to have started before Covid-19, and has not gone back up since restrictions have been removed. Due to Covid and shifting bank holidays there are a few aspects to take into account when assessing the numbers.

Shifted Bank Holidays and Years:

- Christmas Day: 2016, 2021 and 2022
- Boxing Day: 2020 and 2021
- New Year's Day: 2017, 2022 and 2023

Effect of Covid-19 on bank holidays:

- Early May Bank Holiday in 2020 occurred at the end of the first lockdown (1.5 months of full lockdown). This likely contributed to the large amount of incidents that occurred.
- Good Friday and Easter Monday 2020 occurred in the first full lockdown.

Table 1: Covid stringency

Date	Bank Holiday Name	Covid-19 Stringency Index
28/12/2020	Boxing Day	74
28/12/2021	Boxing Day	39
25/12/2020	Christmas Day	73
27/12/2021	Christmas Day	39
08/05/2020	Early May Bank Holiday	80
03/05/2021	Early May Bank Holiday	64
02/05/2022	Early May Bank Holiday	11
13/04/2020	Easter Monday	80
05/04/2021	Easter Monday	73
18/04/2022	Easter Monday	11
10/04/2020	Good Friday	80
02/04/2021	Good Friday	73
15/04/2022	Good Friday	11
01/01/2021	New Year's Day	74
03/01/2022	New Year's Day	39
25/05/2020	Spring bank Holiday	61
31/05/2021	Spring bank Holiday	58
31/08/2020	Summer Bank Holiday	66
30/08/2021	Summer Bank Holiday	23

Assessing Figure 2, it can be seen that:

- On every year but 2022, Boxing Day has around 10% more incidents than Christmas day. 2022 is unusual as both Boxing Day and Christmas day were shifted.
- Christmas Day is normally the quietest bank holiday in the year, followed by New Year's Day. Note that New Year's Day will likely have a lot of backlog of incidents from New Year's Eve to clear.
- Good Friday and Spring Bank Holiday tend to be the busiest bank holidays in the year.
- Bank holidays have a large variation between the busiest and quietest. In 2022, so far, the busiest bank holiday (Good Friday) had 30% more P1-3 incidents than the quietest bank holiday (New Year's Day). This difference has been greater in previous years, with 2021 being 34% and the last year before Covid-19 (2019) being 59%. This demonstrates the need to have different numbers of officers assigned to each bank holiday. The number of officers should align to the numbers of incidents (supply matching demand).

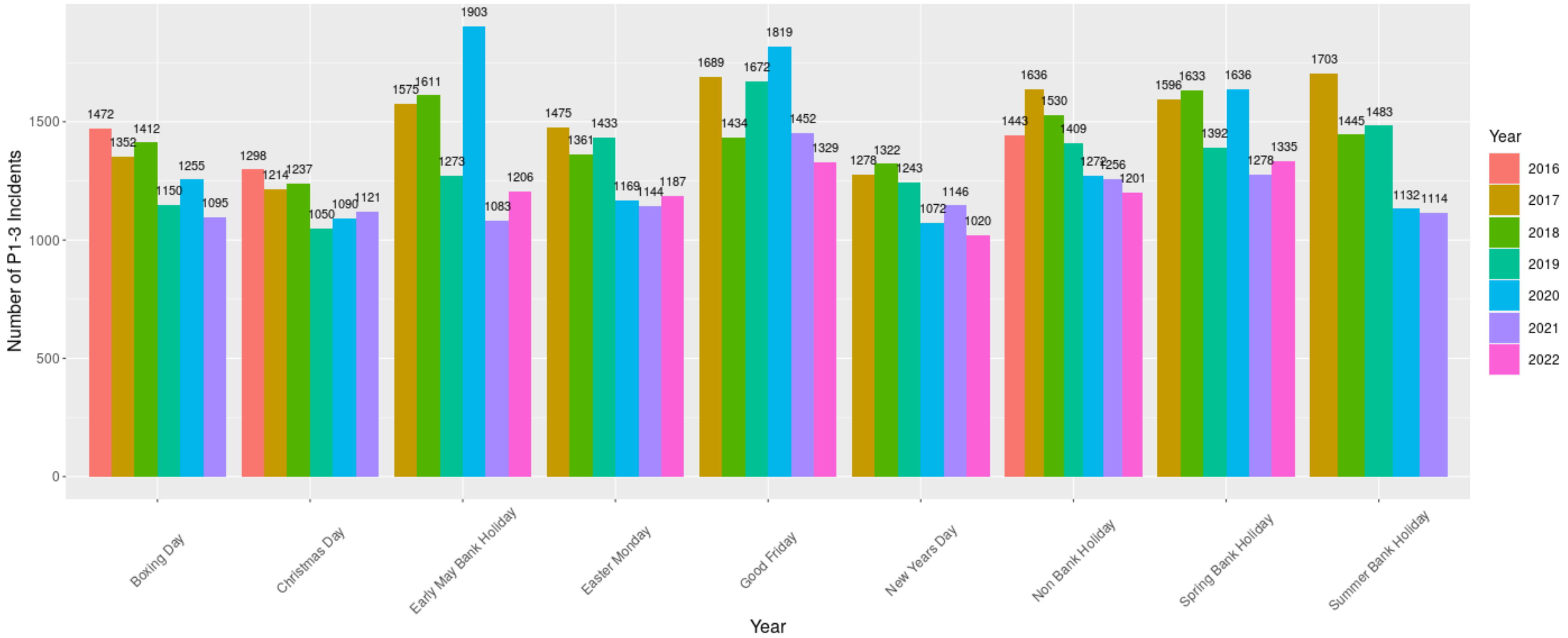

Figure 1: Comparison of Bank Holiday Numbers - Year on Year

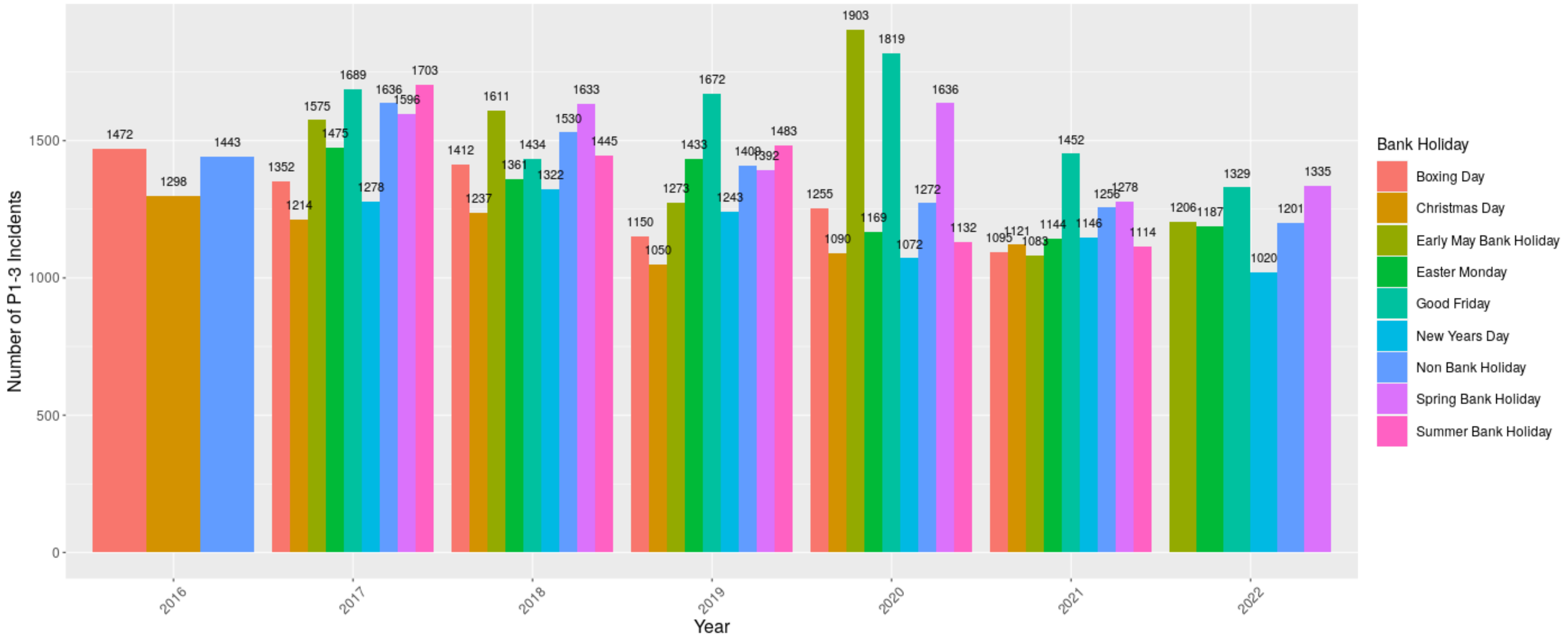

Figure 2: Comparison of Incident Numbers within Year - Bank Holidays

5.2 Assessing the impact at shift and hourly level

Hourly heatmaps were developed to assess the daily peaks and troughs of P1-3 incidents, which can be seen in Appendix A – Mean Hourly Incidents Counts by Bank Holiday. The yearly heatmaps show that, shifting bank holidays (Christmas, Boxing Day and New Year's Day), has an impact on the profiles. This was expected, as even when the bank holiday is shifted, the actual celebration does not (Christmas is celebrated on the 25th if the bank holiday is shifted or not). The tables also show that the general profile of incidents across a bank holiday does not vary year on year (after shifted bank holidays are taken into account).

Simplifying the results down to shift level data, gives table 2, where shifts are defined as; earlies: 7am-3pm, late: 3pm-10pm and nights: 10pm-7am. Assessing the table it can be seen that:

- Christmas Day (when not shifted), has a large proportion of all incidents occurring at night (56.6%), this is much higher than any other bank holiday (next highest was 29.6%, Boxing Day (not shift)).
- When Christmas Day and Boxing Day are shifted, the proportions of P1-3 incidents in each shift, moves back in line with the mean non-bank holiday.
- New Year's Day (non-shifted) has the highest proportion of incidents in the early shift. This is likely due to New Year's Eve night life.

Table 2: P1-3 Incidents - Comparison between shifts across bank holidays

	Mean Number of P1-3 Incidents by Shift			Proportions across Shifts		
Bank Holiday	**Morning**	**Late**	**Night**	**Morning**	**Late**	**Night**
Boxing Day [not shifted]	325	621	398	24.2%	46.2%	29.6%
Boxing Day [shifted]	353	559	261	30.1%	47.7%	22.3%
Christmas Day [not shifted]	300	502	1048	16.2%	27.1%	56.6%
Christmas Day [shifted]	355	583	271	29.4%	48.2%	22.4%
Early May Bank Holiday	390	741	313	27.0%	51.3%	21.7%
Easter Monday	375	663	257	29.0%	51.2%	19.8%
Good Friday	432	738	397	27.6%	47.1%	25.3%
New Year's Day [not shifted]	403	573	221	33.7%	47.9%	18.5%
New Year's Day [shifted]	341	552	255	29.7%	48.1%	22.2%
Non-Bank Holiday	423	665	313	30.2%	47.5%	22.3%
Spring Bank Holiday	401	751	324	27.2%	50.9%	22.0%
Summer Bank Holiday	382	664	327	27.8%	48.4%	23.8%

5.3 Assessing the impact at a priority level

In general, the variation in proportions of P1-3 incidents between bank holidays (and the non-bank holiday) are minimal, as seen in table 3.

Christmas day has the highest proportion of P1 activities (highest priority), but is also one of the days of the year with the lowest number of P1-3 incidents, this shows that the reduction in incidents on Christmas, tends to be within P2 and P3 incidents. This applies to Boxing Day as well, to a lesser extent.

As incident types are related to the priority, it was expected that the proportion of P1-3 incidents would vary between shifts. Assessing the results in table 3, it can be seen that Nights have a higher proportion of P1s than Lates, which have a slightly higher proportion than Earlies for all bank holidays (and the non-bank holiday). The pattern between bank holidays, when split by shift, is similar to that seen in table 3.

Table 4, shows the same data as table 3but with the proportions calculated at day level instead of shift. Again this shows a similar pattern.

Table 3: Mean Proportions of P1-3 across different Bank Holidays

First P1-3 Priority	Boxing Day	Christm as Day	Early May Bank Holiday	Easter Monday	Good Friday	New Year's Day	Non Bank Holiday	Spring Bank Holiday	Summer Bank Holiday
P1	49.6%	53.9%	45.5%	45.6%	46.7%	45.5%	47.3%	47.5%	49.9%
P2	36.1%	33.3%	37.0%	36.7%	38.3%	38.5%	37.2%	36.2%	35.2%
P3	14.3%	12.8%	17.4%	17.7%	15.0%	15.9%	15.5%	16.3%	14.9%
Num Days Included	6	6	6	6	6	6	1968	6	5

Table 4: Mean Proportions of P1-3, by shift, across different Bank Holidays

	Incident First P1-3 Date-time Shift and First P1-3 Priority								
Bank Holiday	**Earlies P1**	**Earlies P2**	**Earlies P3**	**Lates P1**	**Lates P2**	**Lates P3**	**Nights P1**	**Nights P2**	**Nights P3**
Boxing Day	42.2%	38.3%	19.5%	49.6%	36.5%	13.9%	56.6%	33.2%	10.1%
Christmas Day	47.0%	36.8%	16.1%	53.1%	34.3%	12.6%	62.1%	28.1%	9.7%
Early May Bank Holiday	41.4%	39.0%	19.5%	44.6%	37.0%	18.4%	52.9%	34.5%	12.6%
Easter Monday	41.8%	37.7%	20.4%	44.0%	38.0%	18.0%	55.3%	31.8%	12.8%
Good Friday	40.2%	40.5%	19.3%	45.4%	39.1%	15.5%	56.1%	34.5%	9.4%
New Year's Day	42.7%	38.6%	18.7%	44.9%	39.5%	15.6%	51.5%	36.1%	12.4%
Non-Bank Holiday	40.7%	39.4%	19.9%	47.2%	37.6%	15.2%	56.5%	33.2%	10.3%
Spring Bank Holiday	43.1%	38.1%	18.8%	45.9%	36.4%	17.7%	56.7%	33.5%	9.8%
Summer Bank Holiday	44.9%	36.7%	18.4%	50.7%	34.6%	14.7%	54.1%	34.5%	11.5%
Notes:	Shift times are; Earlies: 7am-3pm, Late: 3pm-10pm and Nights: 10pm-7am								

Table 5: Mean Proportions of P1-3 over all shifts, across different Bank Holidays

	Incident First P1-3 Date-time Shift and First P1-3 Priority								
Bank Holiday	**Earlies P1**	**Earlies P2**	**Earlies P3**	**Lates P1**	**Lates P2**	**Lates P3**	**Nights P1**	**Nights P2**	**Nights P3**
Boxing Day	23.1%	17.0%	6.5%	11.0%	9.9%	5.1%	15.5%	9.1%	2.8%
Christmas Day	24.0%	15.5%	5.7%	12.8%	10.0%	4.4%	17.1%	7.7%	2.7%
Early May Bank Holiday	22.9%	19.0%	9.4%	11.2%	10.5%	5.3%	11.4%	7.5%	2.7%
Easter Monday	22.5%	19.5%	9.2%	12.1%	10.9%	5.9%	11.0%	6.3%	2.5%
Good Friday	21.4%	18.4%	7.3%	11.1%	11.2%	5.3%	14.2%	8.7%	2.4%
New Year's Day	21.5%	18.9%	7.5%	13.8%	12.5%	6.0%	10.1%	7.1%	2.4%
Non-Bank Holiday	22.4%	17.8%	7.2%	12.3%	11.9%	6.0%	12.6%	7.4%	2.3%
Spring Bank Holiday	23.4%	18.5%	9.0%	11.7%	10.3%	5.1%	12.5%	7.4%	2.2%
Summer Bank Holiday	24.6%	16.8%	7.1%	12.5%	10.2%	5.1%	12.9%	8.2%	2.7%
Notes:	Shift times are; Earlies: 7am-3pm, Late: 3pm-10pm and Nights: 10pm-7am								

5.4 Forecast Analysis

A time series model was created to understand the impact of bank holidays compared to a general trend, as well as make predictions on future bank holiday demand for adjusting the staffing levels.

A daily time series forecast was undertaken. All incidents were related to shift days (7am-7am) with the total number of P1-3 incidents per shift day modelled. So New Year's Day shift day for 2020 would be; 7am 1st Jan 2020 to the 7am 2nd Jan 2020.

See Appendix B – Model Type and Data Assessment for details.

5.5 Assessing Final Model and Staffing Adjustments

The final model (prophet (multi) (data hols)) test dataset prediction can be seen in Figure 3 with the errors shown in a boxplot in Figure 4. This shows that 75% of estimates were within 7% of the actual result, and 25% of estimates were within 2%. The 1 year forecast for the model can be seen in Figure 5, with the model changepoints demonstrated in Figure 6.

The prophet model has a high level of explainability and is particularly strong at demonstrating seasonal data (which the data shows (weekly and yearly seasonality)). This can be seen in the Figure 7. In this figure it can be seen the general trend in the last 5.5 years has been going down. The weekly trend shows that Friday (7am-7am shift day) is the busiest (10% above week average), followed by Saturday (7.5%), whereas Sunday is the quietest (-6.3%). The yearly seasonality shows a summer demand peak (ranging from -10% in the winter to +10%). The extra regressors (in this case this was only the Covid Stringency index), had minimal impact (+0.5%). The holiday aspect shows peaks or troughs for each bank holiday or national holiday in the next year. From this it can be seen that there are peaks over Halloween/bonfire and for New Year's Eve and troughs for Christmas, Boxing Day and New Year's Day.

Using the model it is possible to predict the next year of P1-3 incidents expected in each shift day. In particular looking at the bank holidays. The result from this can be seen in Figure 8, and are recorded in Table 6. These show that the expected number of P1-3 incidents per shift day varies across the bank holidays in the next year, between 942 (Christmas) and 1319 (Good Friday). The model results show that in general bank holidays do not have a large impact on the expected number of incidents (calls), with half the bank holidays having an effect of less than 2% when compared to the general trend. The bank holiday with the biggest negative impact was Christmas day, which gives a 12.52% reduction compared to the trend.

Using the percentages we can adjust the number of units on bank holidays (using officer / incident ratios and given the number of incidents likely to occur). Currently 138 staff are allocated to each bank holiday. With 8 bank holidays in a year this makes a total of 1104 staff days. The adjusted staff numbers can be found in Table 7. Note that these numbers do not take into account any unactioned incidents from previous shifts (these should be treated as business as usual the next day, but does mean there can be an effect on performance).

Utilising the proportions found in table 2 (Section 5.2), it is possible to obtain estimates for the adjusted numbers of staff at a shift level for each bank holiday. These can be found in table 7.

Table 6: Prophet Estimated Bank Holiday P1-3 Incident Numbers

Bank Holiday	Date	Total Estimate (num P1-3 incidents received in day shift)	Bank Holiday Impact (num incidents) change on trend	Bank Holiday Impact (percent) change on trend	Estimate Proportion of total across all bank holidays
Summer Bank Holiday	29/08/2022	1194	-5	-0.45%	13.22%
Boxing Day	26/12/2022	1085	-13	-1.06%	12.00%
Christmas Day	27/12/2022	942	-150	-12.52%	10.42%
New Year's Day	02/01/2023	960	-104	-8.72%	10.62%
Good Friday	07/04/2023	1319	9	0.75%	14.60%
Easter Monday	10/04/2023	1103	-55	-4.63%	12.20%
Early May Bank Holiday	01/05/2023	1196	43	3.61%	13.24%
Spring Bank Holiday	29/05/2023	1238	22	1.85%	13.70%

Table 7: Adjusted Bank Holiday Staff Numbers – By Shift Day

Bank Holiday	Date	Existing Staff Number	Estimate Proportion of total across all bank holidays	Adjusted Staff Number
Summer Bank Holiday	29/08/2022	138	13.22%	146
Boxing Day	26/12/2022	138	12.00%	133
Christmas Day	27/12/2022	138	10.42%	115
New Year's Day	02/01/2023	138	10.62%	117
Good Friday	07/04/2023	138	14.60%	161
Easter Monday	10/04/2023	138	12.20%	135
Early May Bank Holiday	01/05/2023	138	13.24%	146
Spring Bank Holiday	29/05/2023	138	13.70%	151

Table 8: Adjusted Bank Holiday Staff Numbers – By Shift

Bank Holiday	Shifted	Date	Adjusted Staff Number	Morning	Late	Night	Early Adjusted Staff Number	Late Adjusted Staff Number	Night Adjusted Staff Number
Summer Bank Holiday	NA	29/08/2022	146	27.82%	48.36%	23.82%	41	70	35
Boxing Day	No	26/12/2022	133	24.18%	46.21%	29.61%	32	62	39
Christmas Day	Yes	27/12/2022	115	29.36%	48.22%	22.42%	34	55	26
New Year’s Day	Yes	02/01/2023	117	29.70%	48.08%	22.21%	35	56	26
Good Friday	NA	07/04/2023	161	27.57%	47.10%	25.34%	44	76	41
Easter Monday	NA	10/04/2023	135	28.96%	51.20%	19.85%	39	69	27
Early May Bank Holiday	NA	01/05/2023	146	27.01%	51.32%	21.68%	39	75	32
Spring Bank Holiday	NA	29/05/2023	151	27.17%	50.88%	21.95%	41	77	33

Notes: Shift times are; Earlies: 7am-3pm, Late: 3pm-10pm and Nights: 10pm-7am

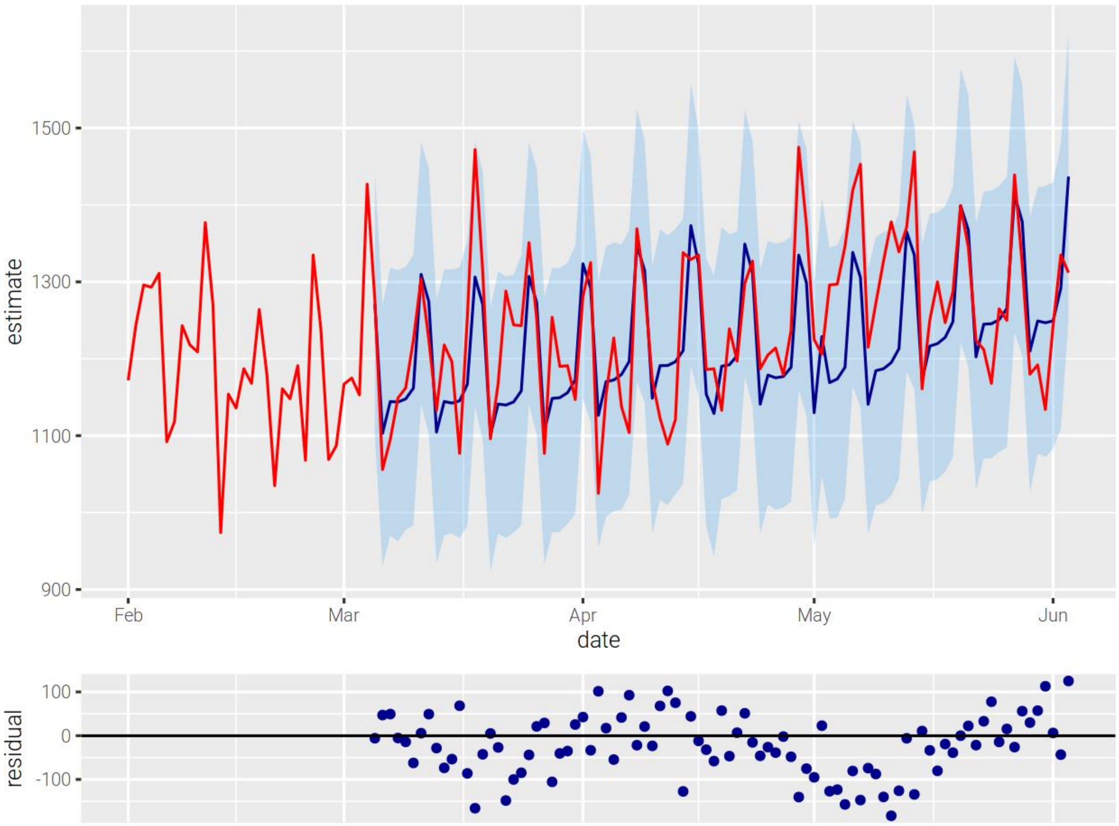

Figure 3: Prophet (multi) (date hols) - Test dataset prediction

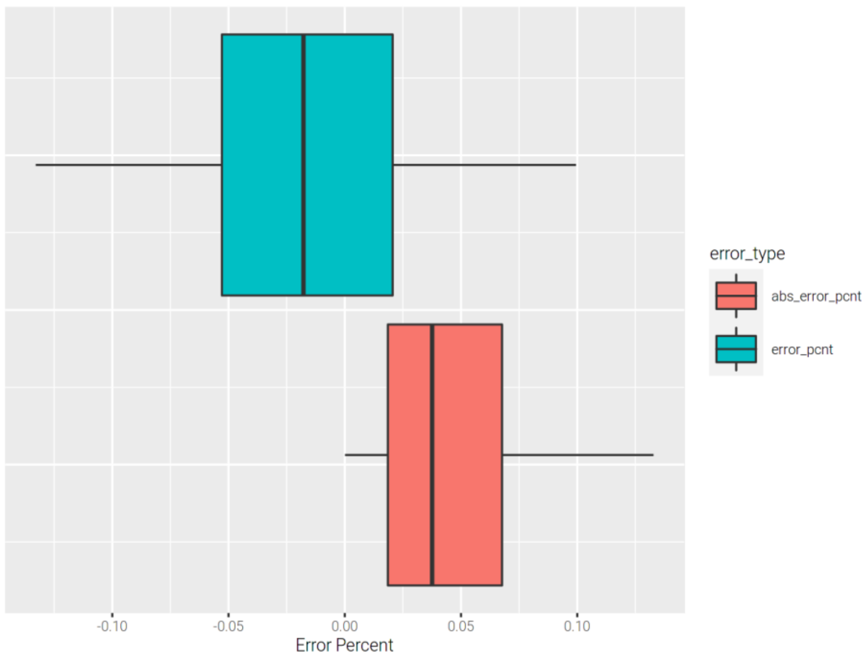

Figure 4: Prophet (multi) (date hols) - Test dataset error boxplot

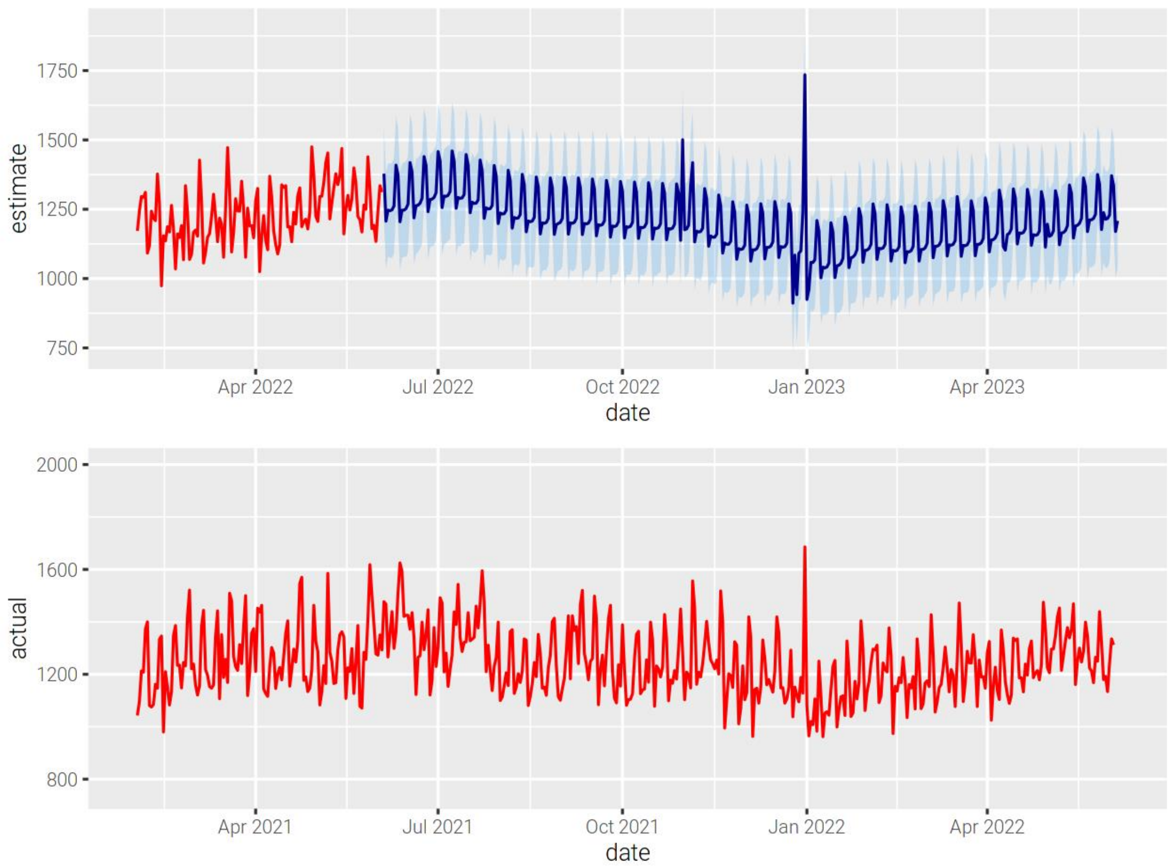

Figure 5: Prophet (multi) (date hols) - One year forecast

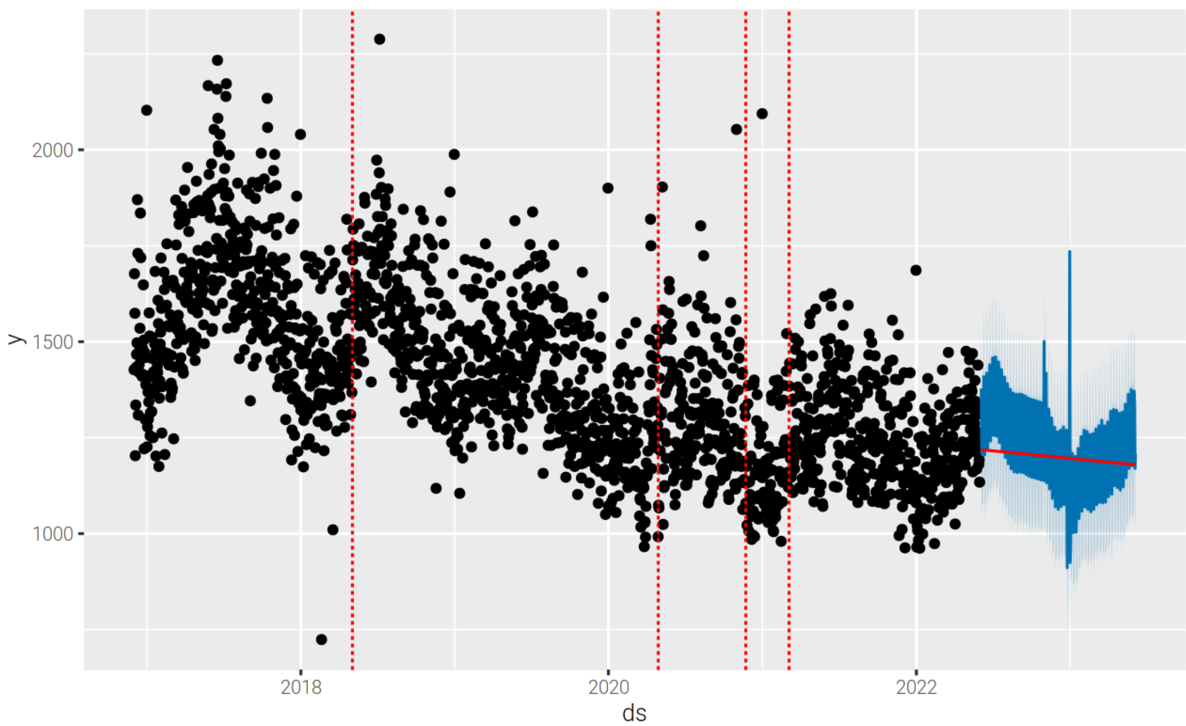

Figure 6: Prophet (multi) (date hols) - One year forecast (with changepoints)

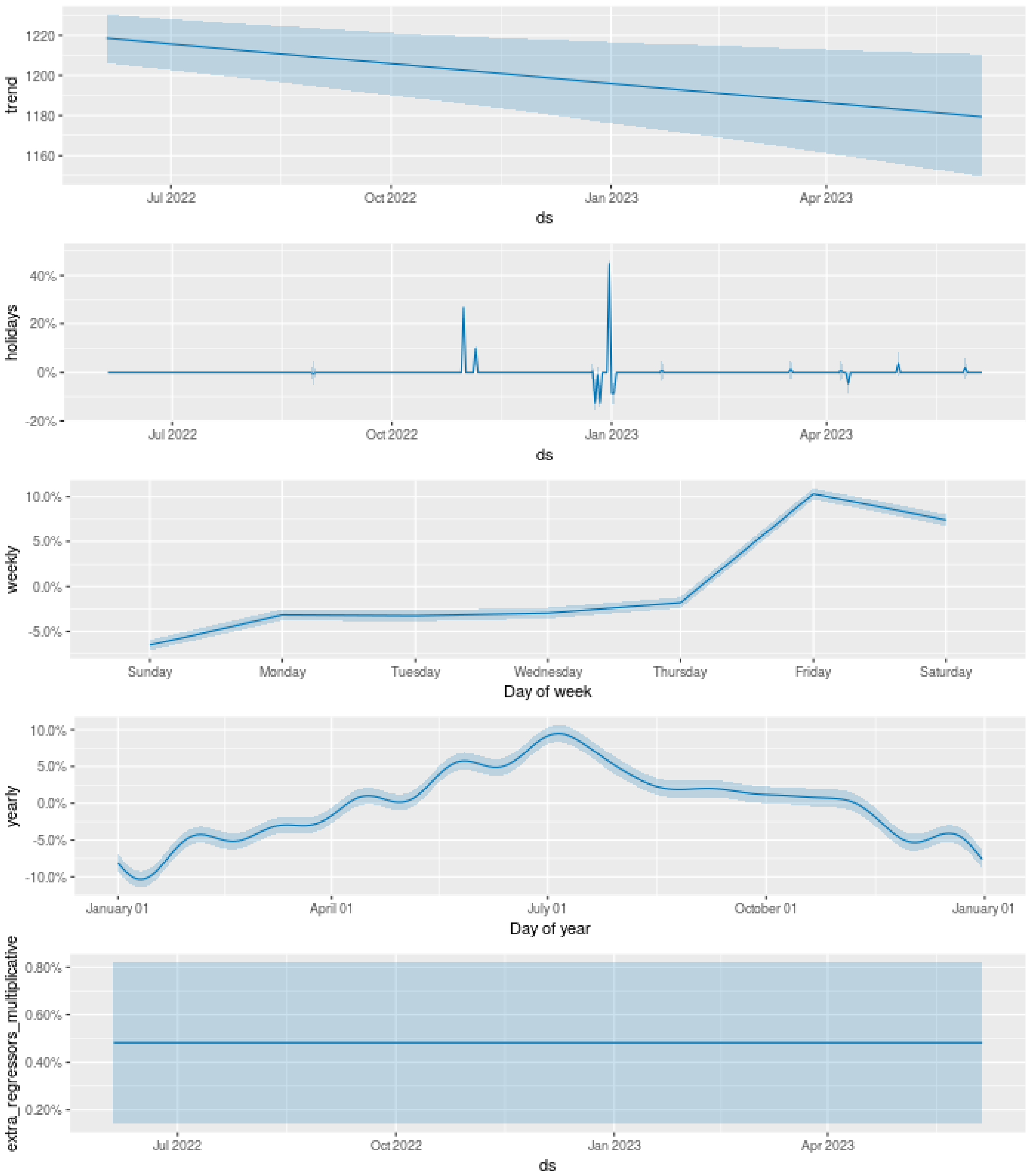

Figure 7: Prophet (multi) (date hols) - Component Plot

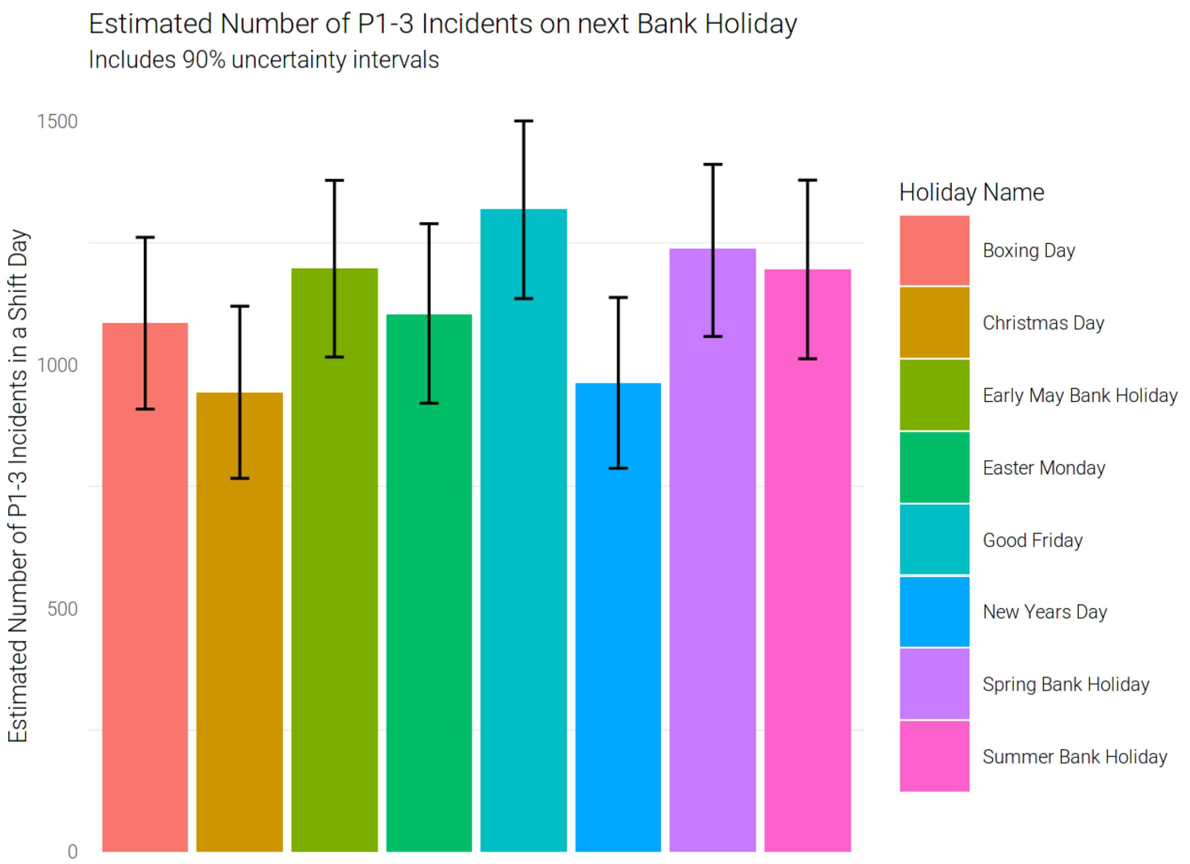

Figure 8: Prophet (multi) (date hols) - Estimated Number of P1-3 Incidents on the next Bank Holiday

Appendix A – Mean Hourly Incidents Counts by Bank Holiday

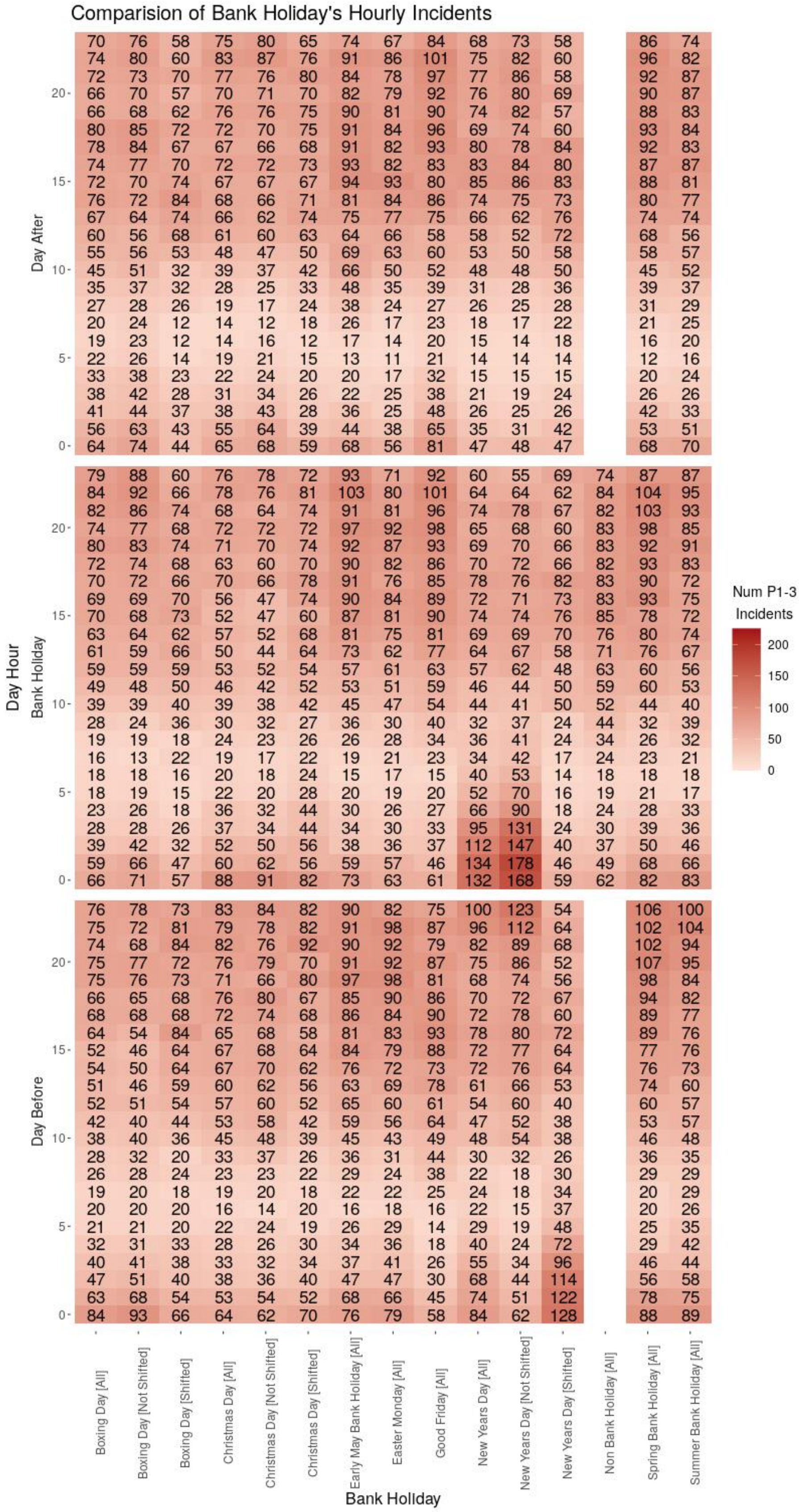

Figure 9: Mean P1-3 Incidents Received per Hour - Bank Holiday Comparison

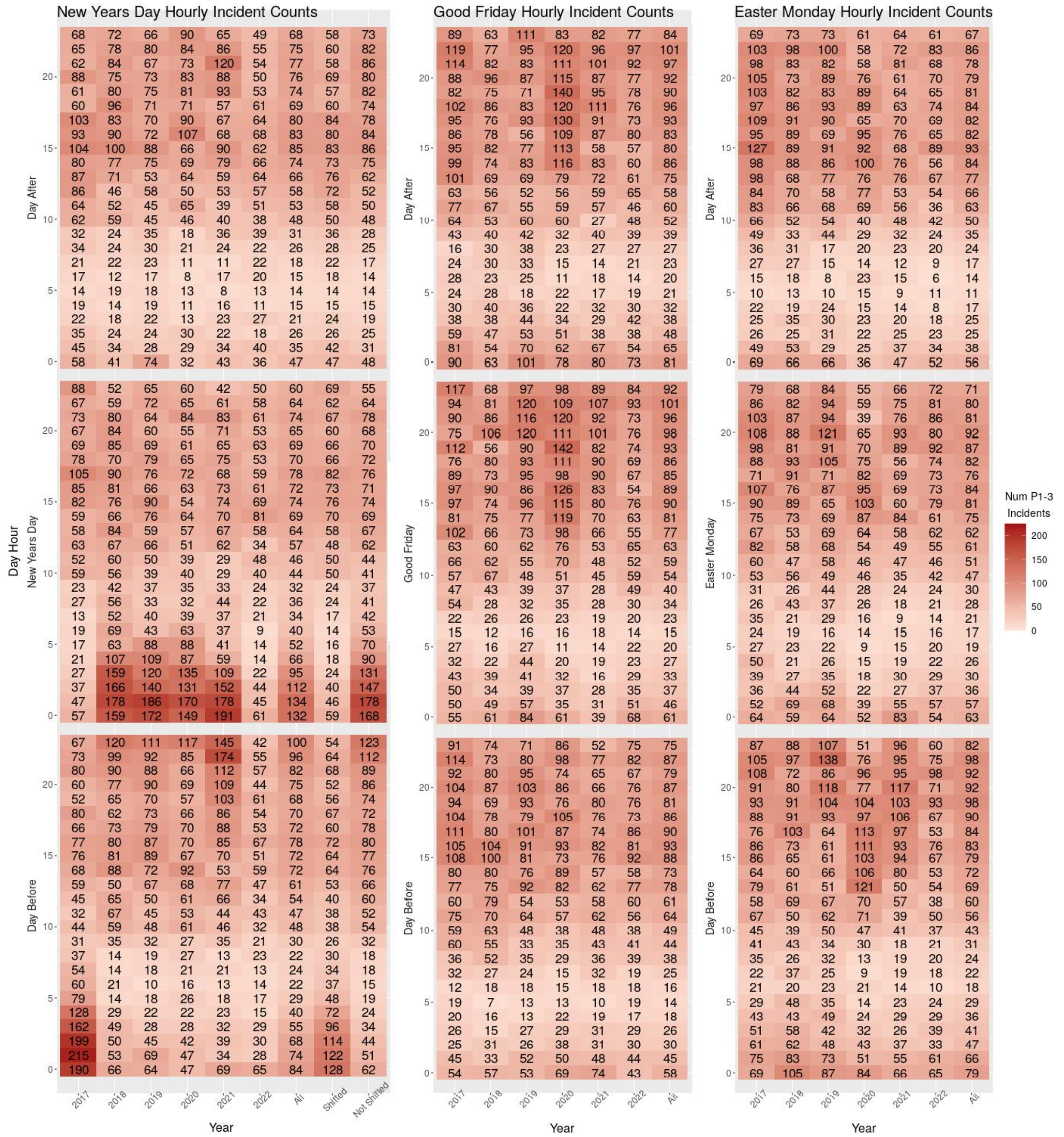

Figure 10: P1-3 Incidents Received per Hour - Year on Year Comparison - New Year, Good Friday and Easter Monday

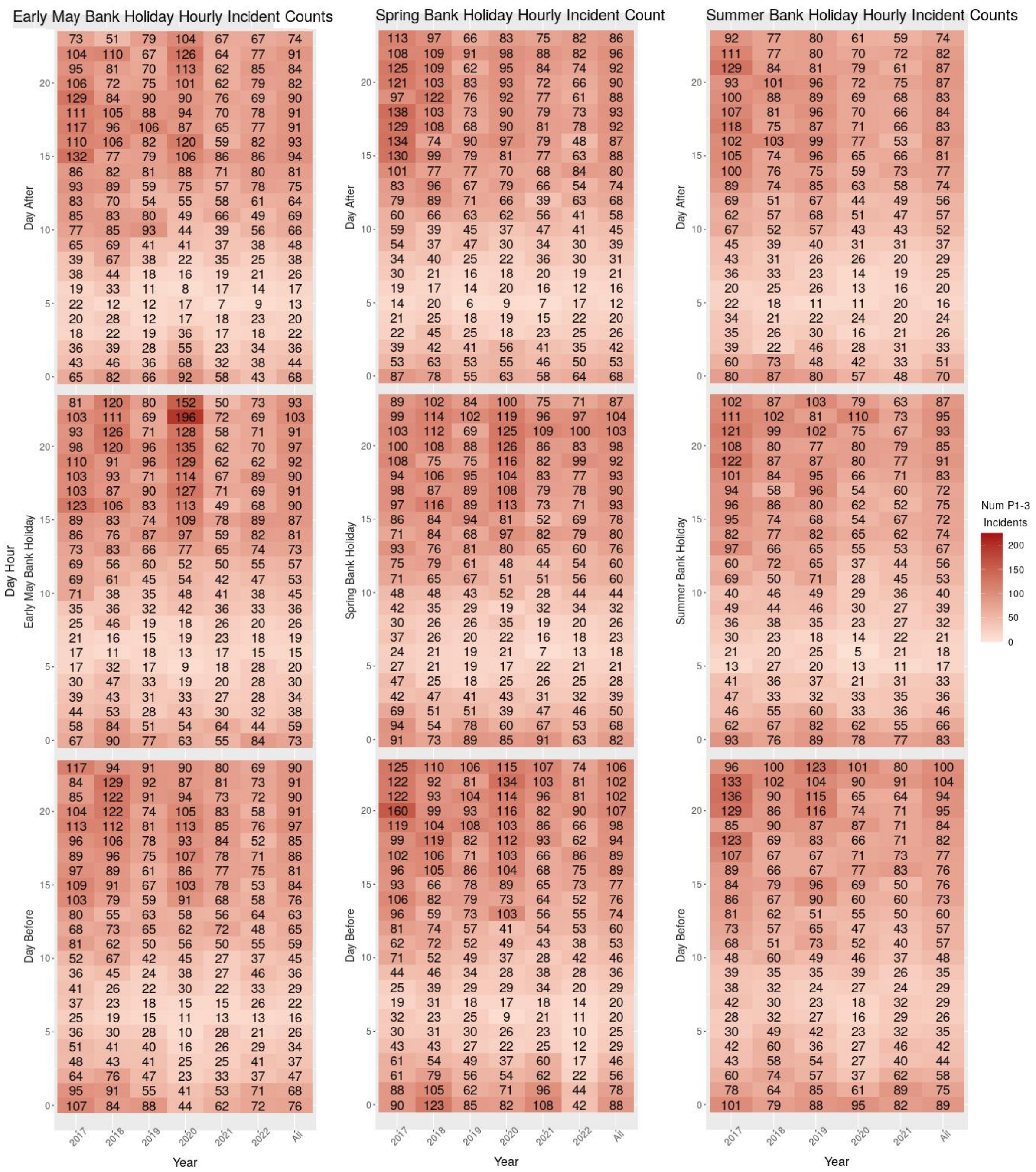

Figure 11: P1-3 Incidents Received per Hour - Year on Year Comparison - Early May, Spring and Summer

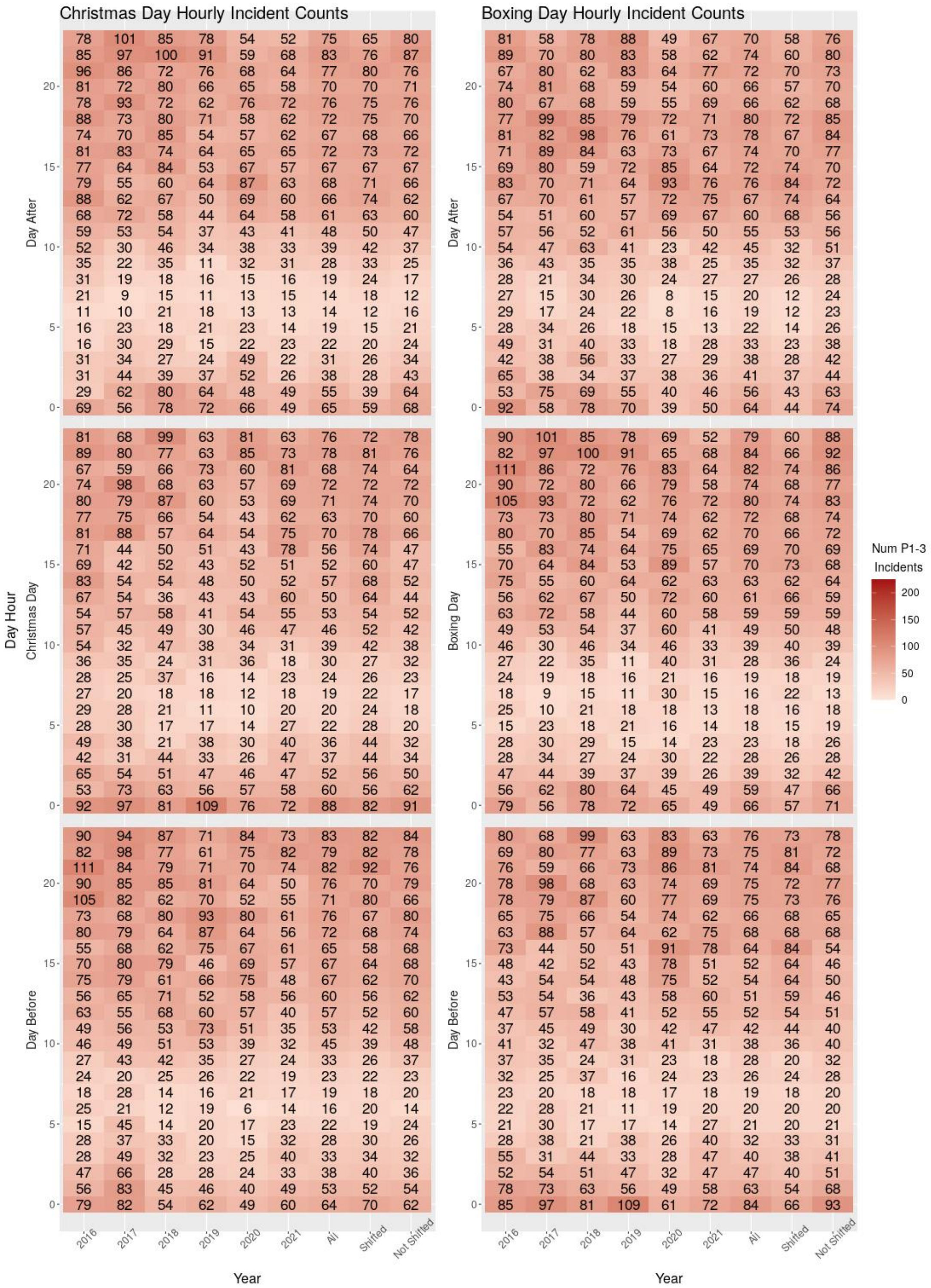

Figure 12: P1-3 Incidents Received per Hour - Year on Year Comparison - Christmas and Boxing Day

Appendix B – Model Type and Data Assessment

Many different models were tested including specialist time series as well as standard regression models using date-parts (instead of dates; year, month, day of week etc. were used as model inputs). The models tested included:

- Time Series Models:
  - Exponential smoothing state space model (ETS)
  - ARIMA
  - Prophet (both additive and multiplicative regressors)
  - Neural Network Time Series (single layer with lags for inputs) (NNAR)
- Regression models using date-parts
  - Random Forest
  - Xgboost
  - MARS
  - Generalized linear model (elasticnet)

All models hyperparameters were optimised using a cross-fold validation (as time series this was a moving window) with five folds used. 91 days (3 months) were removed to be used for testing, see Figure 13. Using the training dataset, the CV plan with 91 day assessment periods was created, Figure 14. A hyperparamter grid size of 500 was used for all the models, except ARIMA where 100 was used due to the time required to calculate. The hyperparameters which gave the smallest mean RMSE across the folds were taken forward.

Multiple combinations of model inputs were also undertaken:

- For Time Series Models:
  - Date Hols
    - Date, Covid stringency index, bank holidays and national celebrations. Note; Bank Holidays are the shifted dates, whereas national holidays include New Year's and Christmas, but they are not shifted. This allows the model to differentiate between shifted and non-shift bank holidays. National celebrations also includes extra holidays such as bonfire night, New Year's Eve and Saint Patrick's Day.
  - Date All
    - Date Hols plus; daylight hours, lunar illumination, football matches (derby's between West Midland's major teams), major protests and televised England football matches in major tournaments.
- For Date-part models:
  - Date-part Hols
    - Year, month, week, day of week, covid stringency index, bank holidays and national celebrations.
  - Date-part All
    - Date-part Hols plus; daylight hours, lunar illumination, football matches, major protests and televised England football matches in major tournaments.
  - Date-part All Extra
    - Date-part All plus; week num, day num in month and num days since 2016.

With the hyperparameters optimised and the best input combination for each model type chosen, the models were fitted on the entire training dataset and assessed using the test dataset (91 days). The goodness of fit statistics can be seen in Table 9. From this table it can be seen that within the test data, XGBOOST (datepart hols) gave the lowest MAPE (4.01%), followed by ARIMA(3,0,3)(1,1,1)[7] (date all) and random forest (datepart hols). Many of the models gave similar predictions, with models of types; xgboost, random forest, arima, prophet and MARS all giving MAPEs between 4-4.7%.

To assess the stability of the fitted models, a moving window approach was used. A combination of 6 folds with assessment periods of 91 days, as well as 3 folds with assessment periods of 365 days was used. These window plans can be seen in Figure 15 and Figure 16. The resultant goodness of fit (GOF) (in the sense of forecasting accuracy) means and standard deviations (sd) are recorded in Table 10. GOF statistics were also calculated on the folds with the long assessment period of 365 days for bank holiday dates only (average of 8 days per year) and recorded in Table 11.

Assessing the results in the Tables (Table 9-Table 11), it can be seen that XGBoost and Prophet are the best model choices. They performed well over the moving windows, and they also gave the most accurate estimates for the bank holiday days. The prophet model was taken forward due to the extra explainability (which allows for ease of calculation of the eventual officer / incident ratios) of the Prophet model (compared to XGBoost (which would require partial plots)) and being a native time series model. Looking at the results for the prophet models, it can be seen that using a multiplicative regressor (instead of additive), resulted in more accurate models.

Comparing the input data choices of “date hols” and “date all”, it can be seen that the additional inputs in “date all” did not improve the model. “date all” achieved better results for the moving window, but worse for the test dataset and bank holiday. Due to this, it was decided to use the “date hols” as it was a simpler model, requiring less additional information to be forecast. So the prophet model (multi) (data hols) was taken forward.

Table 9: Comparing Model GOF's from Testing Dataset

Model Type and Variables Used	MAPE	RMSE	rsq	MAPE Rank	RMSE Rank
PROPHET W/ REGRESSORS (DATE HOLS) (ADD)	4.76	73.99	0.51	10	10
PROPHET W/ REGRESSORS (DATE HOLS) (MULTI)	4.61	72.89	0.52	7	8
PROPHET W/ REGRESSORS (DATE ALL) (ADD)	5.22	80.08	0.52	12	12
PROPHET W/ REGRESSORS (DATE ALL) (MULTI)	4.74	73.73	0.53	8	9
ETS(M,AD,M) (DATE HOLS)	5.55	83.83	0.50	15	15
ETS(M,N,A) (DATE ALL)	5.67	86.46	0.50	16	16
RANDOM_FOREST (DATEPART HOLS)	4.32	66.73	0.57	3	5
RANDOM_FOREST (DATEPART ALL)	5.47	83.57	0.45	14	14
RANDOM_FOREST (DATEPART ALL EXTRA)	4.75	72.28	0.50	9	7
XGBOOST (DATEPART HOLS)	4.01	62.20	0.62	1	1
XGBOOST (DATEPART ALL)	6.77	110.53	0.34	20	20
XGBOOST (DATEPART ALL EXTRA)	4.86	74.71	0.49	11	11
GLM (DATEPART HOLS)	8.50	134.77	0.05	22	22
GLM (DATEPART ALL)	6.14	94.37	0.15	17	17
GLM (DATEPART ALL EXTRA)	6.35	97.75	0.15	18	18

MARS (DATEPART HOLS)	4.39	66.23	0.63	4	3
MARS (DATEPART ALL)	4.43	66.47	0.60	5	4
MARS (DATEPART ALL EXTRA)	4.45	70.53	0.55	6	6
NNAR(1,2,5)[7] (DATE HOLS)	5.35	82.82	0.32	13	13
NNAR(1,1,5)[7] (DATE ALL)	8.05	120.35	0.18	21	21
REGRESSION WITH ARIMA(4,0,1)(2,1,2)[7] (DATE HOLS)	6.66	103.26	0.47	19	19
REGRESSION WITH ARIMA(3,0,3)(1,1,1)[7] (DATE ALL)	4.31	63.82	0.62	2	2

Table 10: Comparing Model GOF Averages and SD's from Moving Windows

Model Type and Variables Used	MAPE mean	MAPE sd	RMSE mean	RMSE sd	rsq mean	rsq sd	MAPE Rank	RMSE Rank
PROPHET W/ REGRESSORS (DATE ALL) (MULTI)	5.44	0.94	84.81	15.78	0.58	0.09	1	1
XGBOOST (DATEPART ALL EXTRA)	5.58	0.73	86.83	14.66	0.55	0.10	2	2
XGBOOST (DATEPART HOLS)	5.73	1.07	88.16	16.49	0.54	0.12	3	3
XGBOOST (DATEPART ALL)	5.81	0.80	92.07	15.03	0.54	0.13	4	7
RANDOM_FOREST (DATEPART HOLS)	5.92	1.14	90.85	16.89	0.52	0.12	5	5
RANDOM_FOREST (DATEPART ALL)	5.93	0.89	90.69	15.30	0.52	0.14	6	4
PROPHET W/ REGRESSORS (DATE HOLS) (MULTI)	5.98	1.68	94.17	28.63	0.57	0.12	7	8
REGRESSION WITH ARIMA(4,0,1)(2,1,2)[7] (DATE HOLS)	5.99	0.79	97.48	16.38	0.54	0.08	8	11
PROPHET W/ REGRESSORS (DATE ALL) (ADD)	6.00	1.32	94.25	20.73	0.58	0.09	9	9
REGRESSION WITH ARIMA(3,0,3)(1,1,1)[7] (DATE ALL)	6.02	1.13	91.96	18.37	0.58	0.06	10	6
MARS (DATEPART ALL)	6.09	1.57	96.94	28.19	0.56	0.14	11	10
PROPHET W/ REGRESSORS (DATE HOLS) (ADD)	6.32	2.08	98.53	31.88	0.56	0.11	12	13
RANDOM_FOREST (DATEPART ALL EXTRA)	6.46	1.45	97.94	19.56	0.50	0.12	13	12
MARS (DATEPART ALL EXTRA)	6.60	1.73	102.97	26.98	0.55	0.09	14	14
MARS (DATEPART HOLS)	6.78	2.00	104.93	31.48	0.56	0.11	15	15
GLM (DATEPART ALL)	7.82	0.64	122.11	12.38	0.19	0.11	16	17
GLM (DATEPART ALL EXTRA)	8.06	0.70	126.07	13.00	0.19	0.11	17	19
GLM (DATEPART HOLS)	8.15	0.70	129.99	15.22	0.11	0.09	18	22
ETS(M,N,A) (DATE ALL)	8.19	3.21	121.17	38.67	0.47	0.12	19	16
NNAR(1,2,5)[7] (DATE HOLS)	8.30	1.77	125.36	24.93	0.17	0.14	20	18
NNAR(1,1,5)[7] (DATE ALL)	8.46	1.21	129.13	17.81	0.16	0.12	21	21
ETS(M,AD,M) (DATE HOLS)	8.82	4.17	128.69	49.92	0.45	0.12	22	20

Table 11: Comparing Model GOF Averages and SD's from 1 Year Assessment Moving Windows (Bank Holidays Only)

Model Type and Variables Used	MAPE mean	MAPE sd	RMSE mean	RMSE sd	rsq mean	rsq sd	MAPE Rank	RMSE Rank
PROPHET W/ REGRESSORS (DATE HOLS) (ADD)	7.70	1.17	108.29	20.17	0.54	0.14	7	8
PROPHET W/ REGRESSORS (DATE HOLS) (MULTI)	6.62	1.88	96.62	21.59	0.55	0.14	2	1
PROPHET W/ REGRESSORS (DATE ALL) (ADD)	7.51	0.52	104.82	2.56	0.51	0.21	6	6
PROPHET W/ REGRESSORS (DATE ALL) (MULTI)	7.19	0.64	100.00	5.84	0.47	0.24	4	4
ETS(M,AD,M) (DATE HOLS)	11.23	3.86	147.55	50.20	0.43	0.33	21	20
ETS(M,N,A) (DATE ALL)	9.85	2.83	130.94	40.14	0.44	0.35	17	17
RANDOM_FOREST (DATEPART HOLS)	9.81	3.99	132.97	43.91	0.69	0.24	16	19
RANDOM_FOREST (DATEPART ALL)	10.04	2.78	132.50	33.26	0.59	0.36	19	18
RANDOM_FOREST (DATEPART ALL EXTRA)	9.86	2.35	126.71	30.58	0.50	0.43	18	16
XGBOOST (DATEPART HOLS)	7.19	3.84	101.78	41.35	0.61	0.31	5	5
XGBOOST (DATEPART ALL)	8.32	2.20	122.83	24.70	0.46	0.39	15	15
XGBOOST (DATEPART ALL EXTRA)	6.42	2.33	97.58	31.93	0.52	0.44	1	2
GLM (DATEPART HOLS)	6.86	2.02	98.45	24.60	0.51	0.12	3	3
GLM (DATEPART ALL)	7.89	3.65	110.57	44.86	0.48	0.25	9	9
GLM (DATEPART ALL EXTRA)	8.03	3.33	112.32	41.61	0.48	0.24	13	12
MARS (DATEPART HOLS)	8.06	1.98	120.32	31.47	0.39	0.40	14	14
MARS (DATEPART ALL)	7.99	2.44	111.35	43.01	0.51	0.42	10	10
MARS (DATEPART ALL EXTRA)	8.00	0.95	108.01	17.69	0.48	0.33	11	7
NNAR(1,2,5)[7] (DATE HOLS)	10.87	1.56	170.01	31.17	0.58	0.11	20	21
NNAR(1,1,5)[7] (DATE ALL)	12.85	1.36	189.84	40.51	0.58	0.23	22	22
REGRESSION WITH ARIMA(4,0,1)(2,1,2)[7] (DATE HOLS)	8.02	0.14	116.05	13.57	0.33	0.16	12	13
REGRESSION WITH ARIMA(3,0,3)(1,1,1)[7] (DATE ALL)	7.84	0.20	111.88	1.74	0.54	0.08	8	11

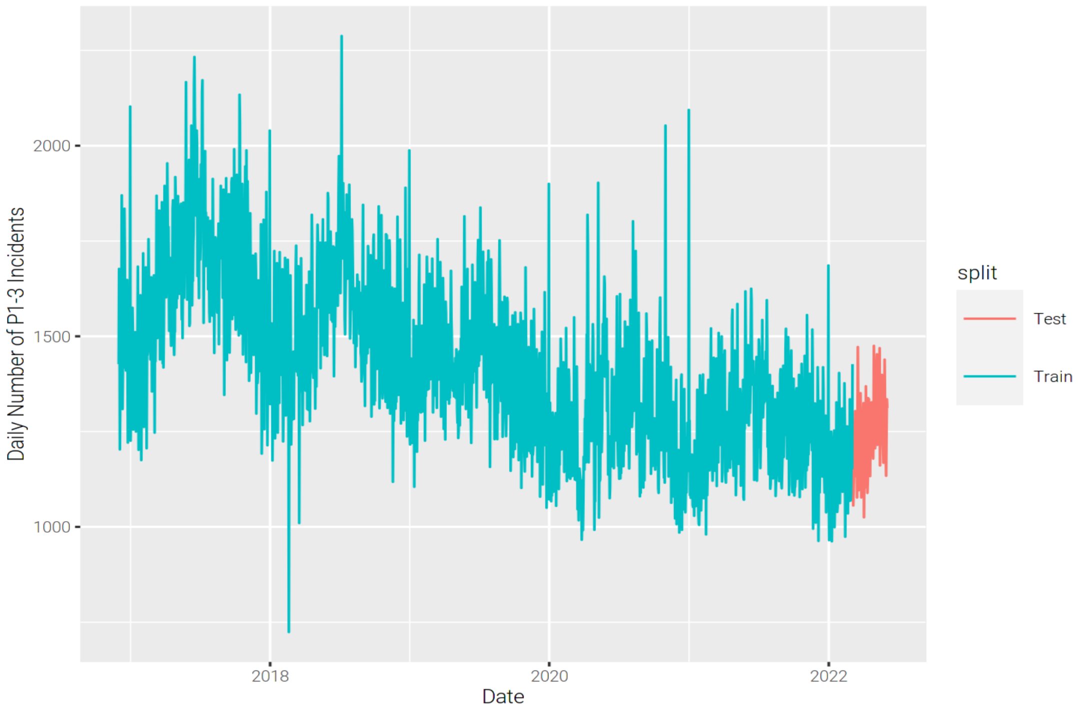

Figure 13: Time Series Testing/Training Split

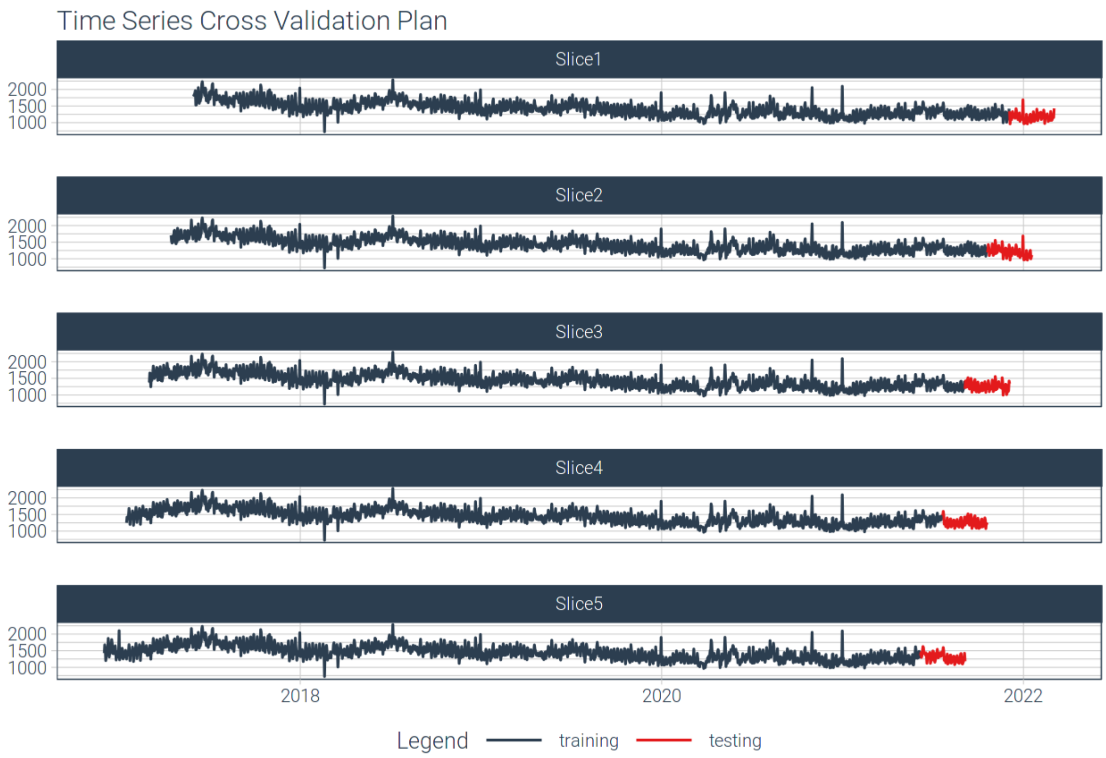

Figure 14: Time Series CV Plan

Figure 15: Resample (moving window) Plan (short, 91 day assessment)

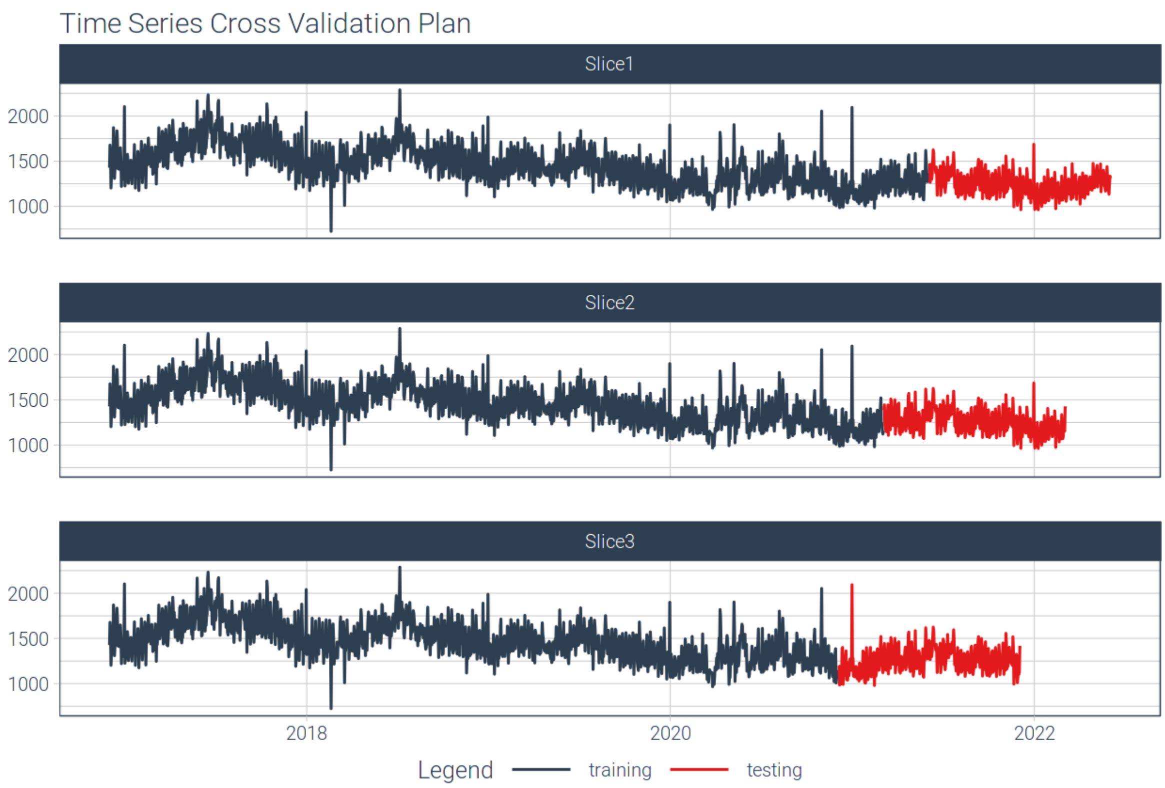

Figure 16: Resample (moving window) Plan (long, 365 day assessment)