Evaluating the Impact of the New Triage Policy on Emergency Room Activity

Pedro Miguel Sousa Magalhães

This study evaluates the impact of the new mandatory telephone pre-triage system on the number of emergency room visits. To assess this impact, publicly available emergency statistics from the Vila do Conde emergency room, one of the three hospitals involved in the trial, were used. Based on insights from domain experts, the emergency room in Barcelos served as a control group. The study focuses exclusively on the use of SARIMA models and transfer functions to develop a simple counterfactual as a quick means of identifying potential areas for further investigation. The findings suggest that, although the intervention led to a reduction in emergency occurrences during the period from January 2024 to September 2024, the impact was modest. The study concludes with several additional questions that could help refine the analysis and motivate further research, potentially using more advanced modeling techniques

1. Introduction

1.1 Motivation

The high pressure on emergency rooms has long been a significant concern for the Portuguese National Health System. This issue becomes particularly problematic during periods of increased demand, such as the winter season, when the pressure is exacerbated by the seasonal flu and other similar illnesses. In response to these challenges, public health authorities introduced a mandatory telephone pre-triage system, which was initially trialed in December 2023 at a limited number of locations. Alongside this, a new policy was implemented to refer patients to private hospitals. The mandatory pre-triage system is now promoted as a central component of the government's strategy to alleviate emergency room congestion.

This study is motivated by an initial analysis conducted by Professor Pedro Pita Barros, an economics professor at Nova SBE and a recognized expert in health economics. His initial findings, published as a blog on December 18, 2024, examined emergency occurrences at the Unidade de Saúde de Vila do Conde (ULS Vila do Conde), one of the trial locations, and tracked the evolution of emergency room visits following the introduction of the mandatory pre-triage system in January 2024. In his study, the city of Barcelos was used as a comparison group to assess the potential impact of the intervention (see Barros and SBE 2024b). **From this initial work he concludes that there evidence to support that the introduction of this new policy lead to a 10% drop in emergencies.**

This study goal is to use traditional time series statistics to study the same data and compare if the results of a model driven analysis are aligned with the initial findings suggested by (Barros and SBE 2024b). For this ARMA models will be used and later expanded using dynamic time series models to incorporate the effect of a reference series. **This study approach focus on using a counter-factual to simulate a series without any intervention and later compare to real measurements.** The goal is not to challenge or further support the 10% drop claim, instead, given the number of assumptions taken, is to conclude if indeed this intervention had any significant impact.

This analysis uses the exact same assumptions as (Barros and SBE 2024b) original post, namely the use of total emergencies as the relevant indicator and the assumption of ULS Barcelos as a reference series. In the conclusions some additional questions are risen regarding impacts and assumptions, namely if total emergencies is most appropriate indicator.

1.2 Exploring data on emergencies for Vila do Conde and Barcelos

This study focuses exclusively on data from two institutions:

- **Hospital Santa Maria Maior, EPE**, which was renamed in 2024 to **Unidade Local de Saúde de Barcelos**.
- **Centro Hospitalar Póvoa de Varzim/Vila do Conde, EPE**, which was renamed in 2024 to **Unidade Local de Saúde de Póvoa de Varzim/Vila do Conde**.

The dataset includes information on various types of emergency incidents. However, this analysis focuses specifically on general emergencies. The data was retrieved from the public portal of the Portuguese National Health Service at https://transparencia.sns.gov.pt/explore/dataset/atendimentos-por-tipo-de-urgencia-hospitalar-link/information/?sort=tempo.

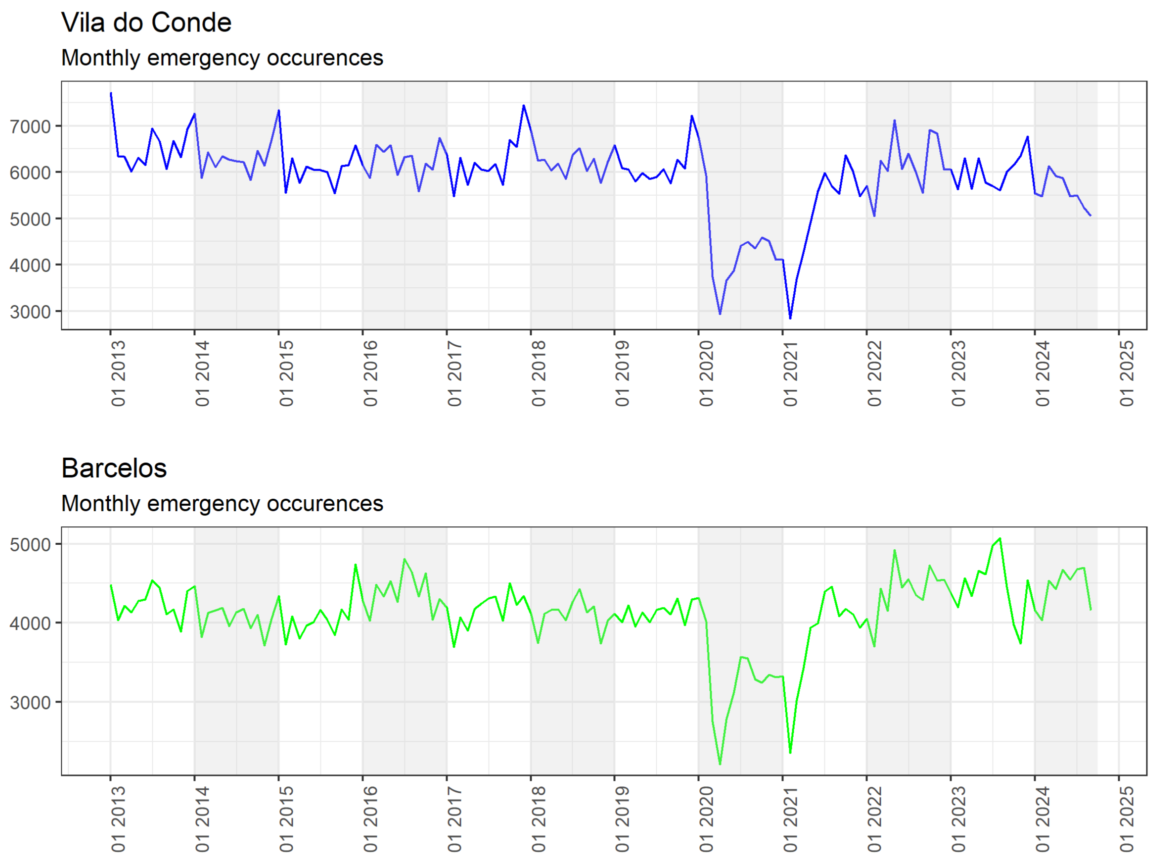
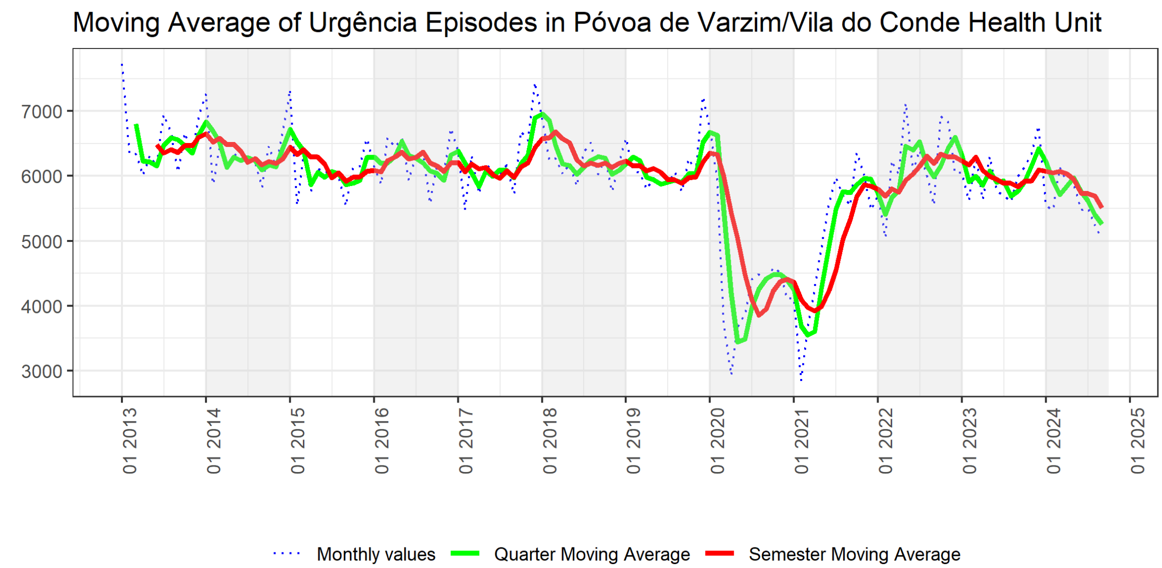


Figure 1: Monthly emergencies for both intervened and control series

There is a significant period of outlier behavior between January 2020 and April 2024, corresponding with the COVID-19 pandemic. This period caused substantial disruptions, which must be addressed before further modeling. The extended duration and magnitude of these anomalies require a strategy for outlier compensation.

Seasonality: The data exhibits signs of a **12-month seasonal pattern**, with local peaks typically occurring toward the end of each year.

Trends: Although no clear overarching trend is evident, there is a noticeable **continuous decrease in emergency occurrences** in Vila do Conde during the most recent months.



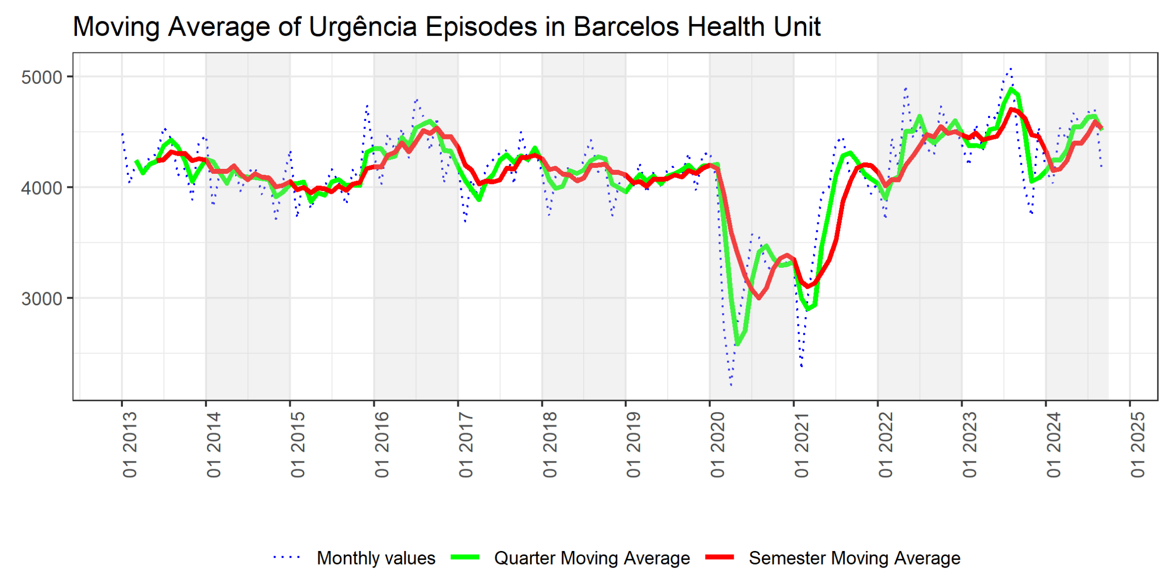


Figure 2: Monthly emergencies smoothed by Quarter and Semester moving average

An initial examination of the monthly emergency data for both locations revealed the following insights:

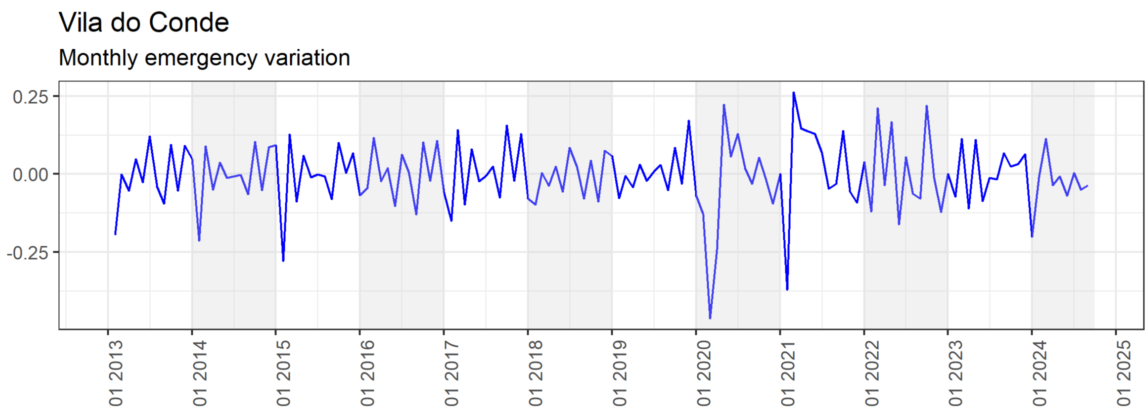
Smoothing both series using 3-month and 6-month moving averages highlights the following:

- The significant impact of the COVID-19 pandemic, which seems to have had lingering effects, as the series after that period show changes in trend and patterns.

- The pre-COVID series exhibited a very tenuous trend. Despite some periods, such as 2016 in Barcelos and 2018 in Vila do Conde, initial visual inspection reveals no clear significant change in the series mean. These periods appear to be point-in-time interventions or outliers. Analyzing each of these is beyond the scope of this study.
- The post-COVID period shows notable changes compared to the previous period, suggesting that the pandemic left lingering effects over time. In Vila do Conde, an upward trend in 2012 seems to have been followed by a downward trend starting in 2013 (the mandatory pre-triage intervention began in January 2024). A quarterly moving average below the semiannual average implies that this trend persists over time. In contrast, Barcelos has shown a continuous upward trend since the end of the pandemic. The sharp decline at the end of 2023 appears to be an outlier due to a shortage of practitioners, leading to the closure of several emergency services during that period. In 2024, Barcelos seems to demonstrate a sustainable upward trend.

A major takeaway from the visual inspection is that both time series are **non-stationary**. This conclusion is further supported by the **Kwiatkowski–Phillips–Schmidt–Shin (KPSS)** and **Augmented Dickey–Fuller (ADF)** tests. Since the goal is to study the impact of an intervention or policy change on the rate of increase or decrease in emergency visits, transforming the main series into a **logarithmic rate of change** allows us to achieve a similar goal while simplifying the modeling process. Let \$X\$ represent the random variable of monthly emergencies, and we will use the following variable for our analysis.

$$Z_n = ln(\frac{X_n}{X_{n-1}})$$



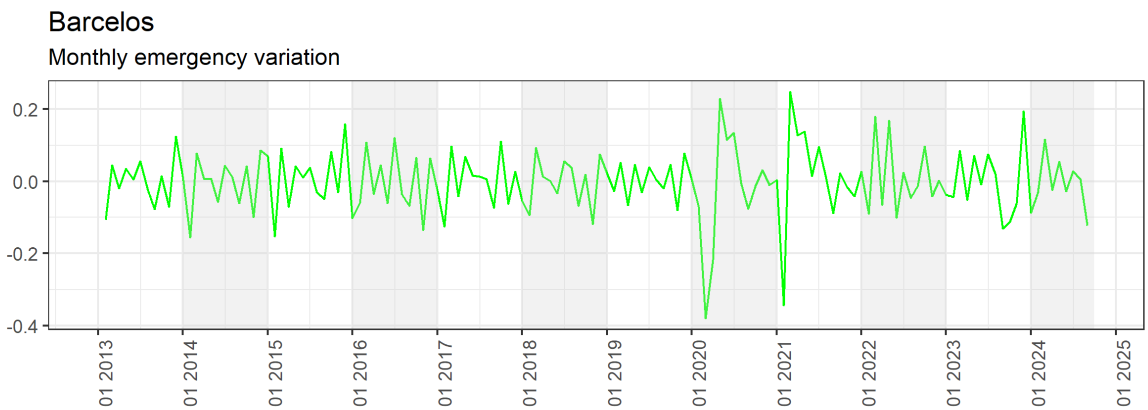
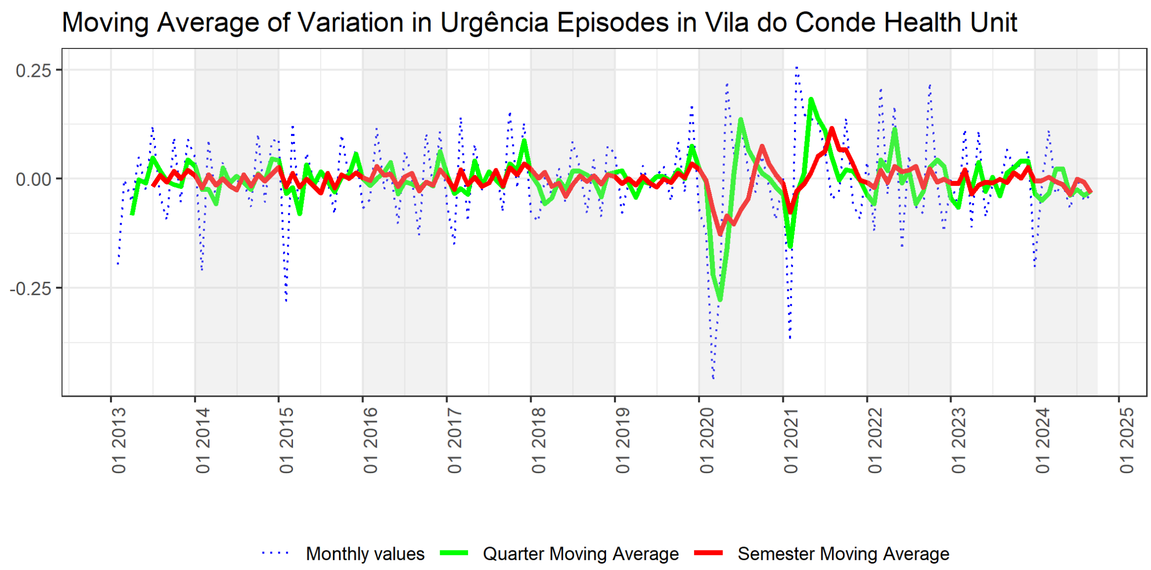


Figure 3: Monthly emergencies Log variations



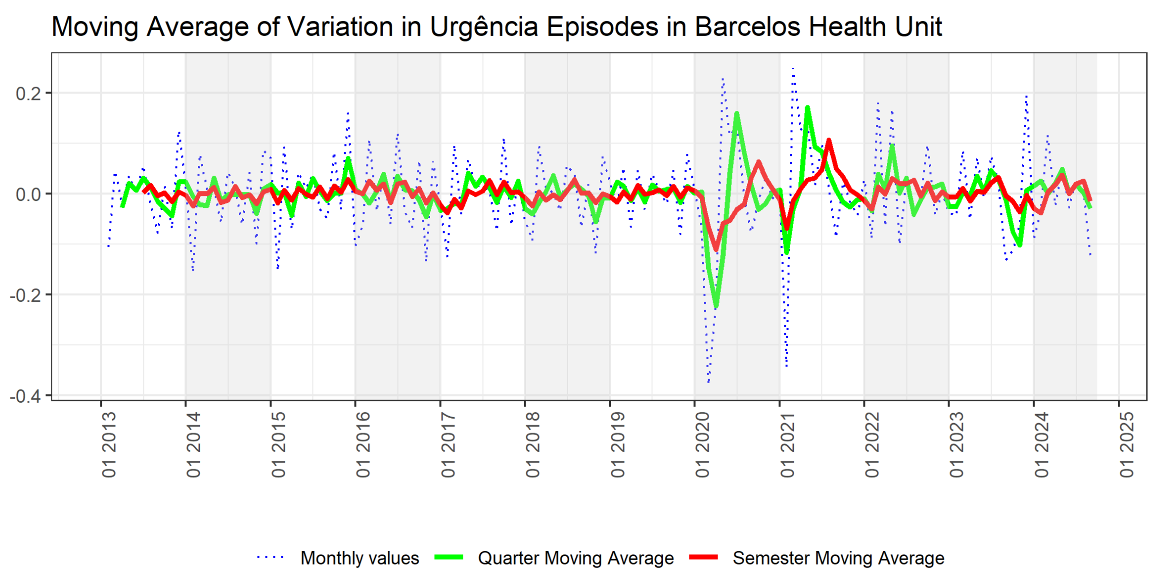


Figure 4: Monthly emergencies Log variations and moving average

Fig show the transformed series with stationary despite the clear outlier moments during covid.

2. Assessing the impact of new pre-triage policy

2.1 Approach

Intervention analysis, introduced by Box and Tiao (1975), provides a framework for evaluating the effect of an intervention on a time series under study. The approach assumes that the intervention affects the process by altering the mean function or trend of the time series (Cryer and Chan 2009a , pp. 249).

In this context, addressing the mandatory pre-triage system as a new intervention is a valid approach. Given the nature of the intervention, it is necessary to model not only the short-term or immediate effects but also any spillover effects, as such changes are expected to generate lingering impacts over time with variable effects. A simple step or pulse function would not adequately capture the full scope of the intervention's impact. However, due to limited domain knowledge, it is not feasible to estimate all potential effects precisely.

Barros' article offers valuable insights by presenting Barcelos as a related series, with shared characteristics that make it a suitable control group for comparison (see Barros and SBE 2024b). It is important to note, however, that using another series to measure the intervention's impact comes with certain caveats:

- **Local interventions in Barcelos:** Emergency services in Barcelos were subject to local interventions, such as decisions made by hospitals, which could influence the analysis. It will be assumed that these local interventions did not significantly affect the outcome of the study.
- **Outliers in both series:** As previously noted, both series contain outliers, which may be caused by events such as strikes or doctor shortages. For the purposes of this study, these periods will not be excluded or treated as outliers (e.g., using a pulse dummy variable and modeling it as an external variable). Therefore, it is assumed that, while these outliers exist, they will not substantially impact the modeling process or the final results. This is a simplification, and addressing these outliers would require additional analysis outside the scope of the current study.

2.1.1 Covid-19

Both the original and transformed (log returns) series show a significant disruption during the COVID-19 pandemic period. The available metadata is insufficient to determine whether emergencies during this period were recorded in a separate dataset or if these two institutions were closed or operating under limited capacity. However, unlike other periods of identified outliers in both series, the magnitude and duration of the COVID-19 impact are substantial enough to significantly affect any model generated from this time series.

An intervention or outlier approach using external variables could be useful in this case(Cryer and Chan 2009a , pp. 257). However, the primary objective here is to use historical data to infer the existence of a significant impact from a recent policy change. Given these circumstances, the focus is on identifying strong trends and seasonality, rather than analyzing specific past effects, which, while influential, are not replicable despite any spillover effects they may have had. Therefore, the COVID-19 impact will be "removed" from the series by "forecasting" the data between January 2020 and March 2021.[1]

2.2 Methodology used to assess the intervention

Given the proposed goal and the assumptions outlined in the previous analysis, the impact of the mandatory telephone pre-triage (hereafter referred to as **SNS24 triage**) at ULS Vila do Conde will be evaluated as follows:

1. A counterfactual series will be forecasted using historical data from Vila do Conde, with **Barcelos** serving as an external variable and acting as a control group.
2. Data from 2013 to 2020 will be used to replace the COVID-19 period from January 2020 to March 2021.
3. The newly generated series for the period 2013-2023 will be used to fit a model, which will then generate the counterfactual. Given the objectives of this study, no smoothing will be applied between the actual series from 2013 to 2020 and the COVID-19 replacement period.

3. Counterfactual and measuring intervention

3.1. Covid-19 outlier period

Figures 4 and 5 isolate the series between February 2013 and December 2019 (the original series includes January 2013, but the first observation was removed due to the transformation to avoid 0 or infinite variations). As discussed earlier, both series show signs of outliers, particularly around the start of 2014 and 2015.

In order to remove the unwanted effects of the covid period an ARMA model will be used to reconstruct both ULS Vila do Conde and ULS Barcelos series. The series used transformed to log variation and both adf and kpss test support the assumption of stationarity for the series untill January 2020 as seen on the table below.

[1] It's important to notice that forecasting in order to reconstruct a period of a time series does introduce a substantial level of bias and noise to the analysis. Although it's assumed that this effect does not impact the outcome and conclusions, further analysis is required to validate this assumption in future works.

```
# A tibble: 6 x 4
# Groups:   unidade_saude [2]
  unidade_saude test      adf.returns.test kpss.returns.test
  <chr>         <chr>                <dbl>             <dbl>
1 Barcelos      statistic            -4.80            0.0481
2 Barcelos      p.value               0.01            0.1
3 Barcelos      parameter             4               3
4 Vila do Conde statistic            -5.73            0.151
5 Vila do Conde p.value               0.01            0.1
6 Vila do Conde parameter             4               3
```

The **Autocorrelation Function (ACF)** and **Partial Autocorrelation Function (PACF)** show substantial correlation at 12 and 24 lags in both series, suggesting a yearly pattern, consistent with monthly data. This aligns with expectations given the nature of the dataset.

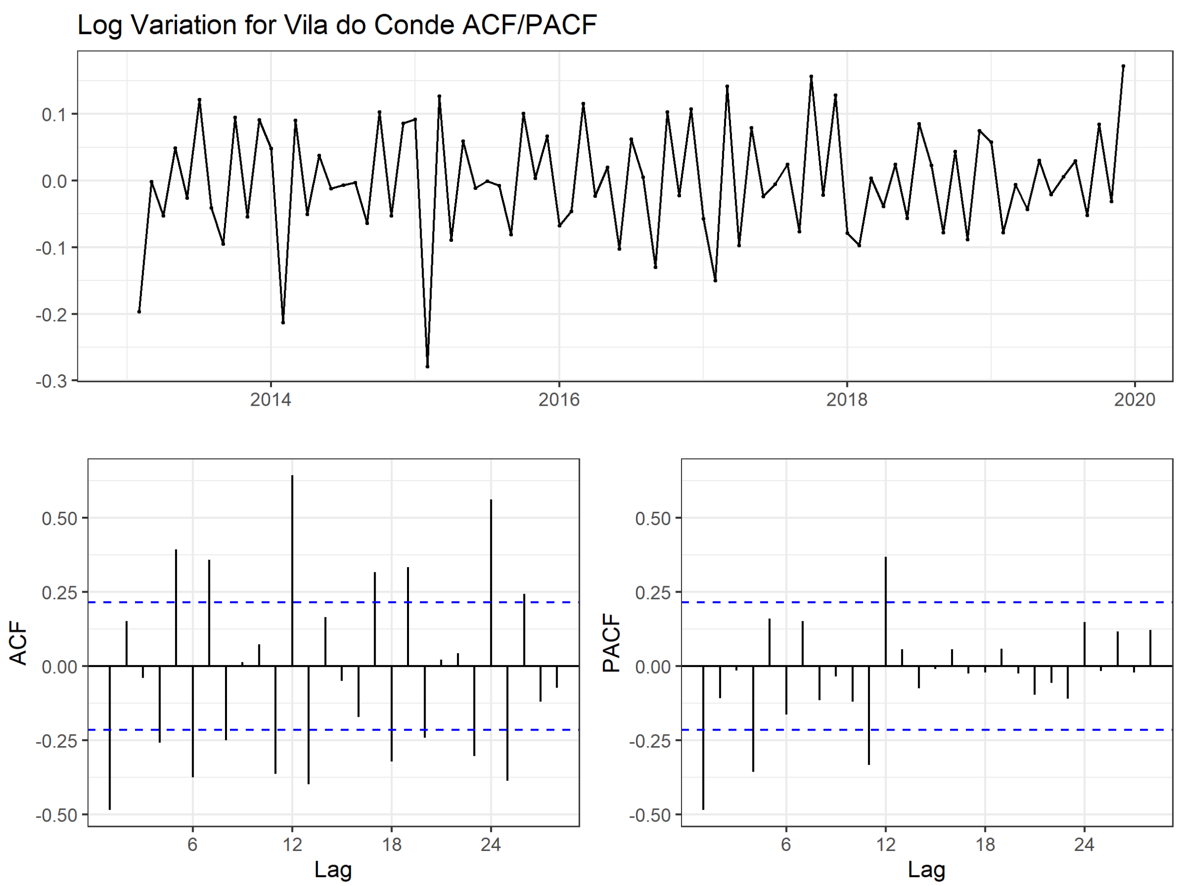


Figure 5: ACF and PACF for Vila do Conde log variation

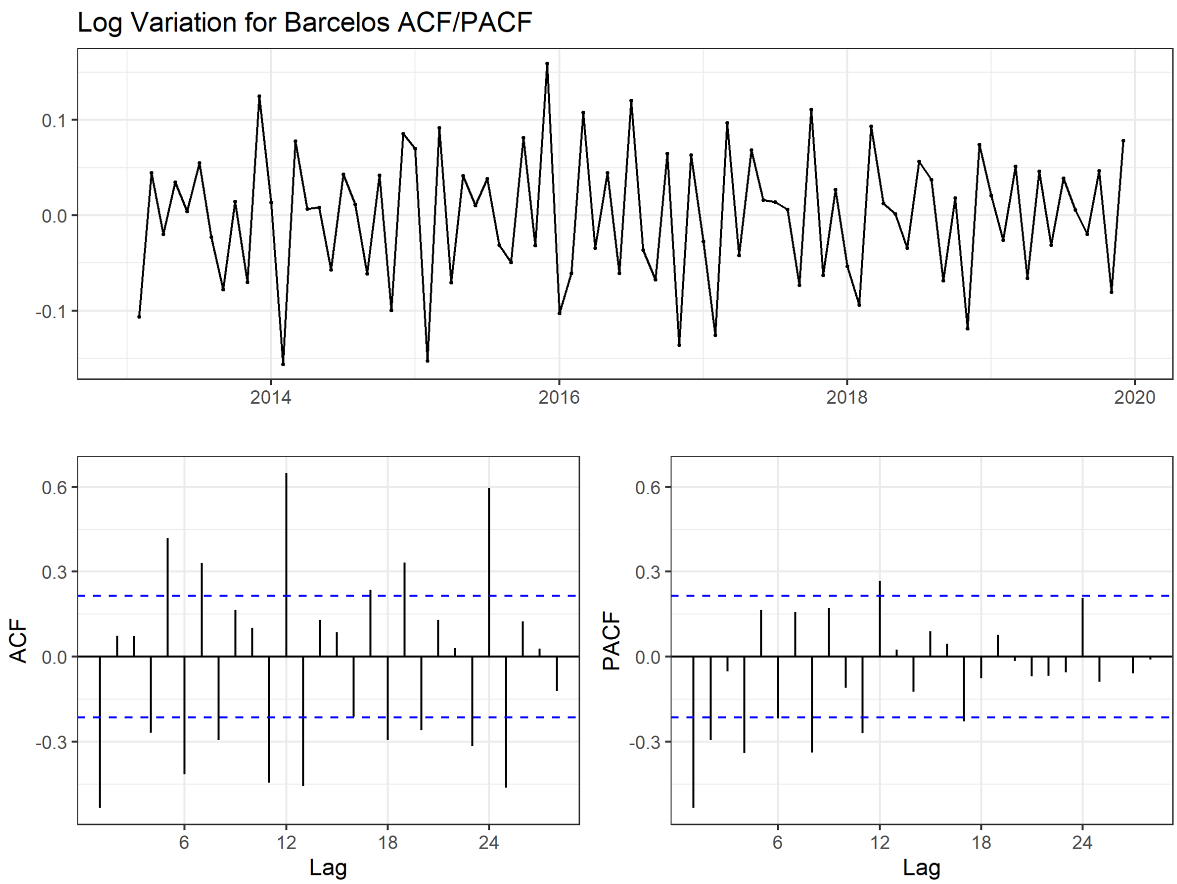


Figure 6: ACF and PACF for Barcelos log variation

Significant lags also exist at the 6 and 18 lags suggesting multi sesonality. Intuition suggests that a 6 month cycle could exist given the number of emergencies due to seasonal illnesses like flu which typical are at its high during winter.

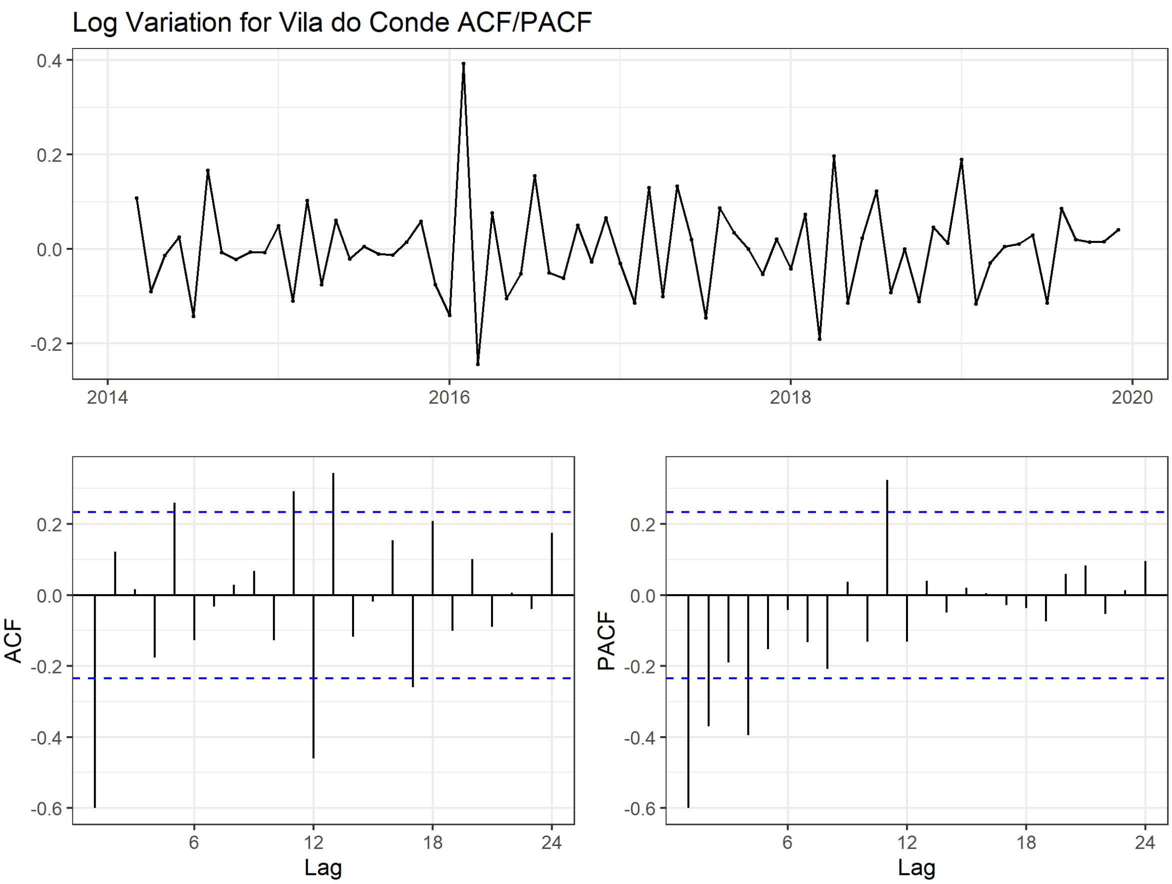


Figure 7: ACF and PACF for Vila do Conde log variation for (1-B)12

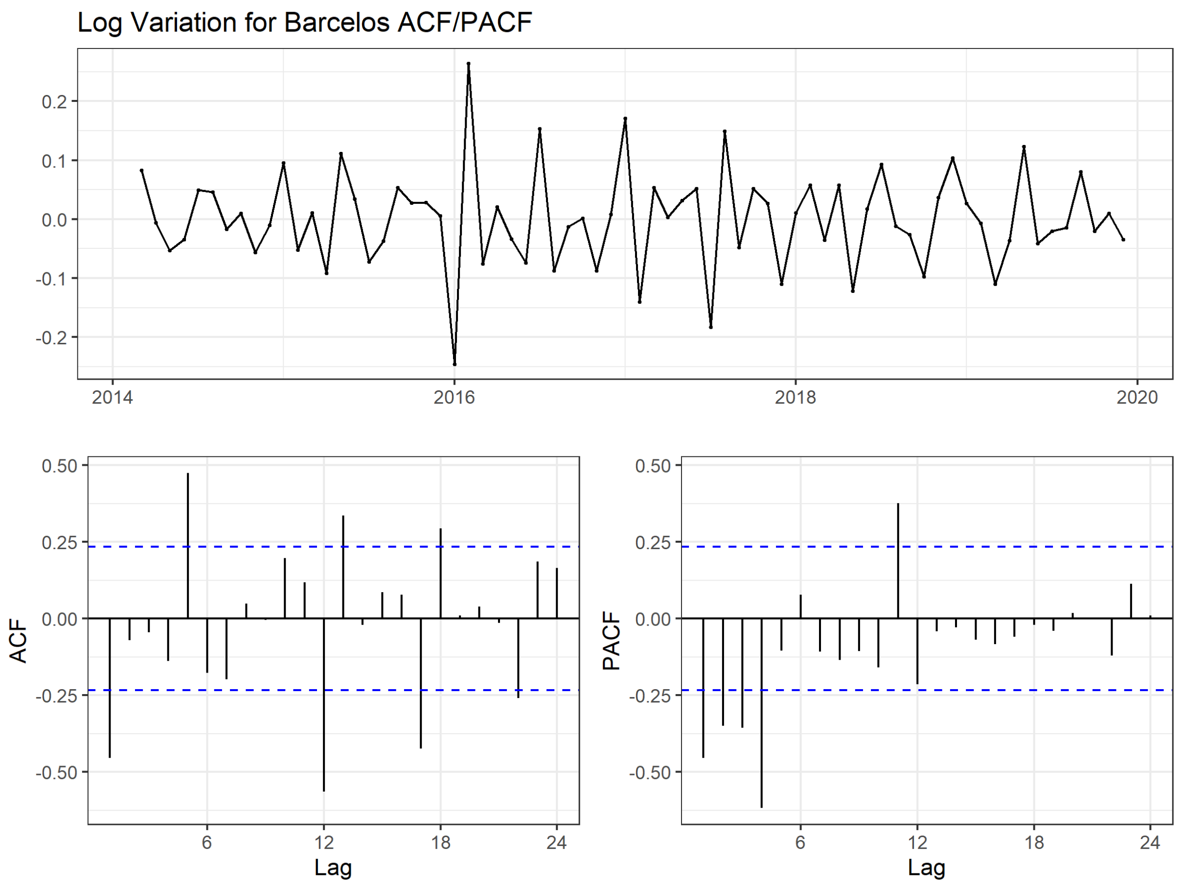


Figure 8: ACF and PACF for Barcelos log variation for (1-B)12

Analyzing both ACF/PACF and residuals a SARIMA(1,0,1) x (0,1,1)12 was used to model the pre-covid moment for Vila do Conde and a SARIMA(4,0,1) x (1,1,1)12 for Barcelos. Below the model statistics support this models as good approximations for the reality under study. [2]

[2]Further detail about pre-covid modeling can be found on appendix B.

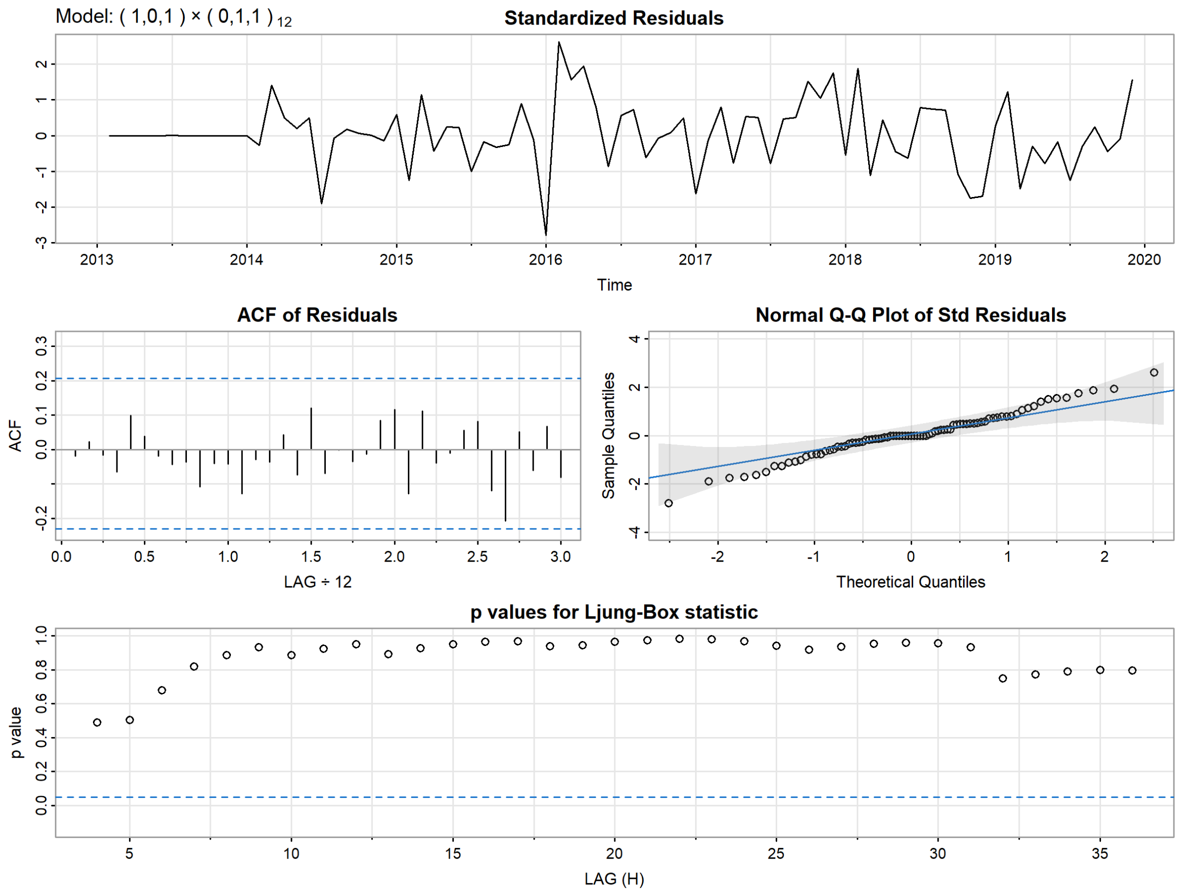


Figure 9: Vila do Conde Pre-covid model fit statistics

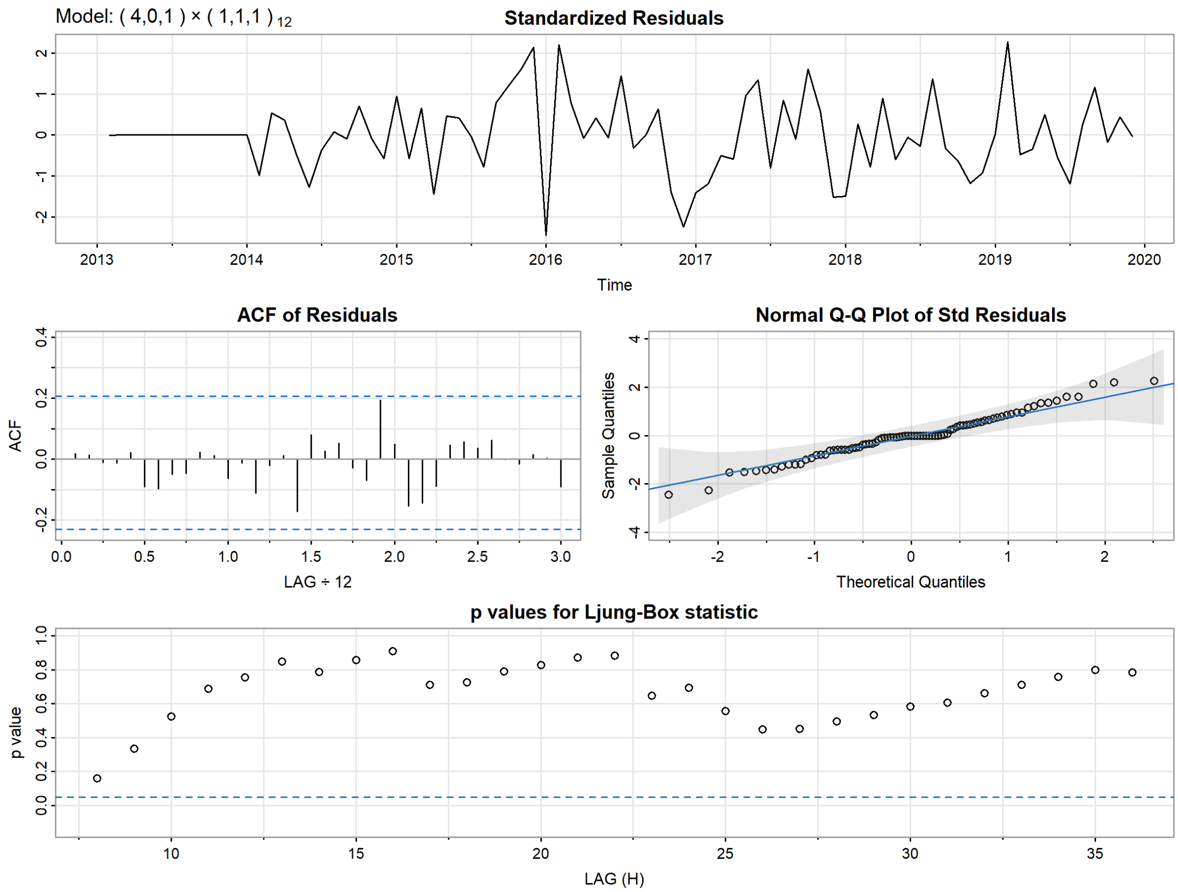
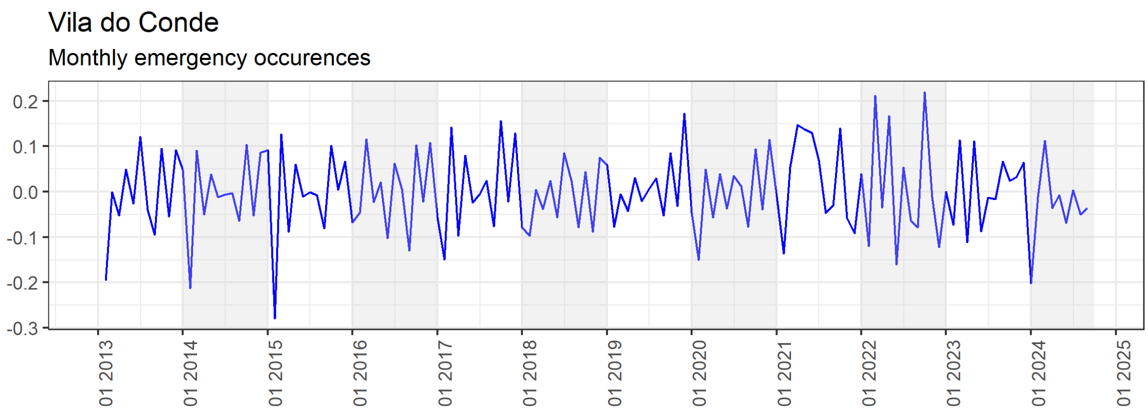


Figure 10: Barcelos Pre-covid model fit statistics

Using this model its now possible to replace the covid effect based on previous periods:



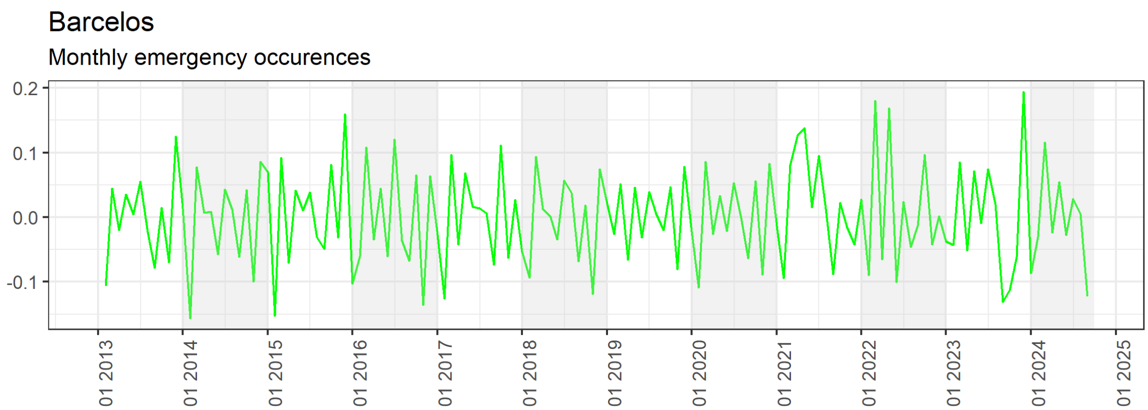


Figure 11: Series without Covid effect for both Barcelos and Vila do Conde

3.2. Fitting a Counterfactual

After correcting the series from 2013 to September 2024, it is necessary to verify whether the initial hypothesis, supported by domain knowledge, that **Barcelos** reflects a similar reality and can thus serve as a valid control series, holds true. Cross-correlation between both series shows a strong correlation at lag 0, indicating a strong instantaneous effect. This confirms that Barcelos can be considered a suitable candidate for the control group.

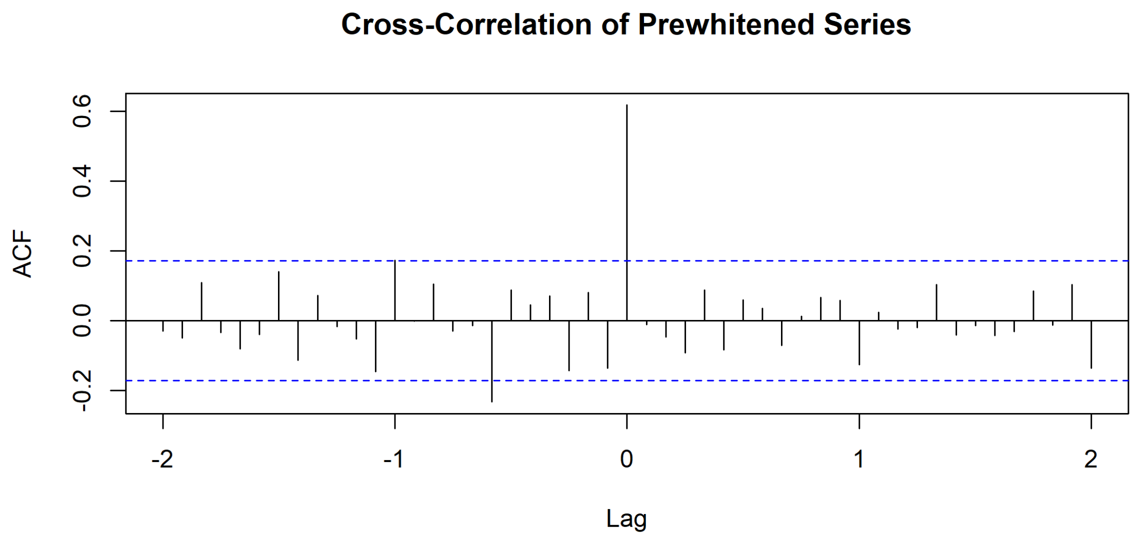


Figure 12: CCF between Barcelos and Vila do Conde

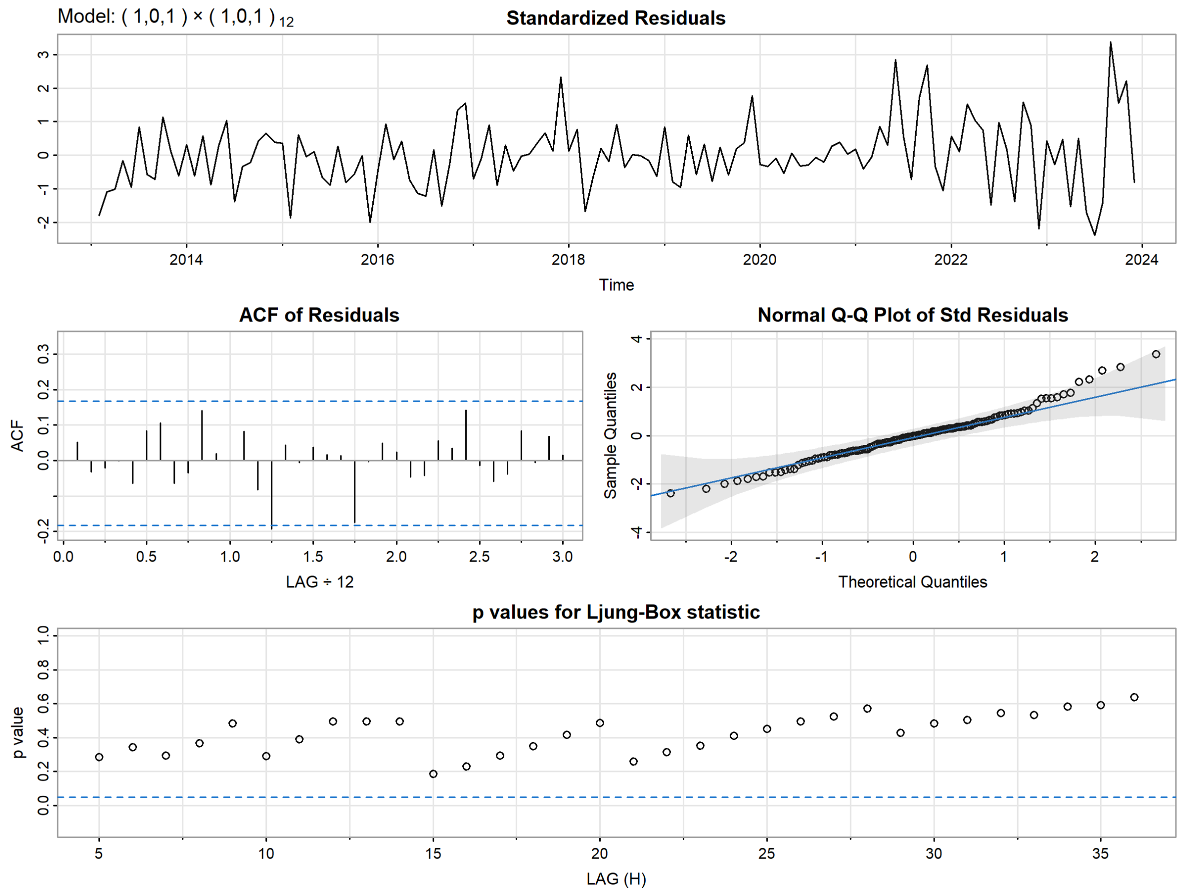


Figure 13: Counter factual model fit

Fitting a **SARIMA (1,0,1) x (1,0,1)** model with a 12-month seasonality and using **ULS Barcelos** as an exogenous variable produces a "good enough" model for the purposes of this analysis. The residuals show low correlation, suggesting that the model is adequately fitting the data. This model can then generate a "theoretical" series for the period between January 2024 and September 2024, the only available period during the intervention.

Visual inspection of **Figure 14** shows that both the sample data and the counterfactual closely follow one another. A **Welch two-sample t-test** confirms what can intuitively be inferred from the plot: both series are likely derived from the same distribution, as the null hypothesis (H0: difference in means = 0) cannot be rejected.

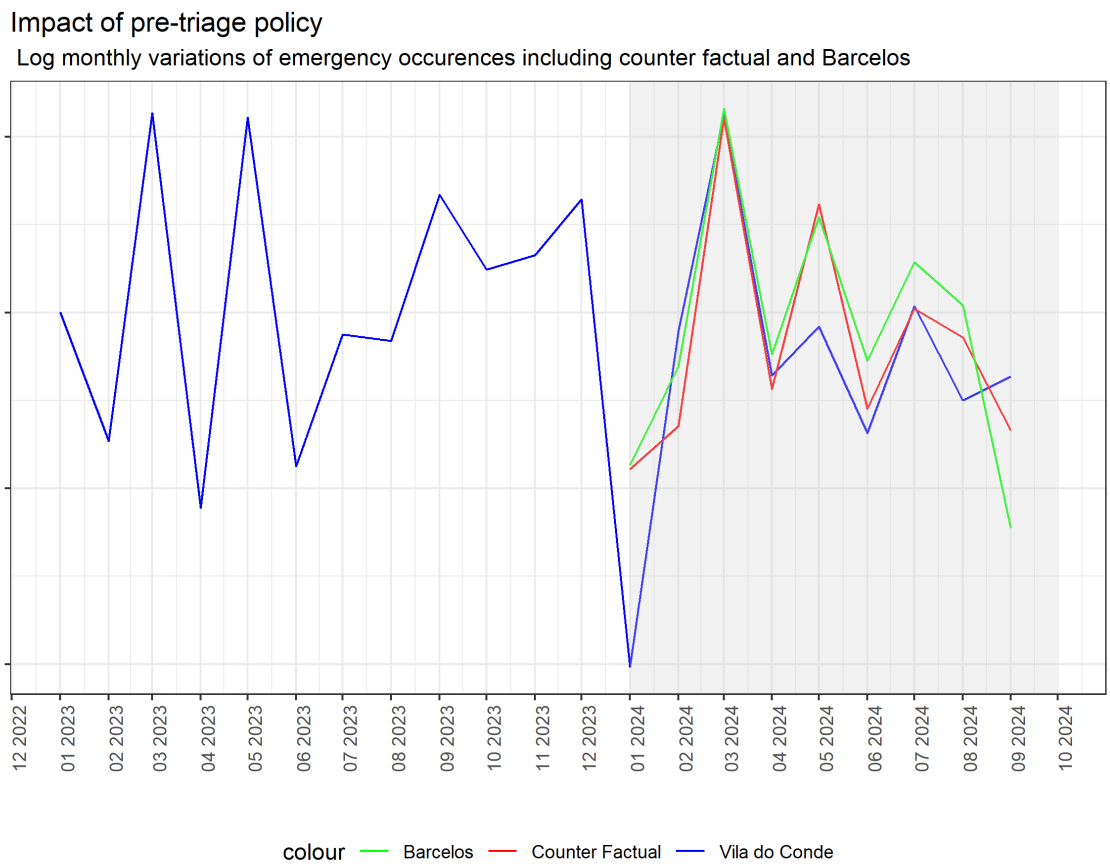


Figure 14: Comparing impact

```
Adding missing grouping variables: `unidade_saude`
```

```
# A tibble: 1 x 2
  unidade_saude `mean(diff)`
  <chr>                <dbl>
1 Vila do Conde        -1.38
```

The average difference between counter factual and real measurements if of -1.38 percentage points. Meaning, that the introduction of this policy lead to an increase reduction of 1.38 percentage points. This is aligned with the visual inspection of figure 14 where the counter factual follows rather closely the real series with significant differences at 01/2024 (lack of smoothing?) and 05/2024.

4. Conclusion and future work

By modeling a counter-factual model, it was possible to show that the introduction of this new policy lead to an additional reduction in number of emergencies in ULS Vila Conde, which is aligned with Barros and SBE (2024b) initial remarks. Nonetheless, the impact seem to be far less pronounced raising the question if it could be consider substantial enough to justify the nation wide adoption. Despite the found effect, the model shows evidence that the great part of the behavior observer during the trial period would occurred despite the introduction of this new policy.

Since the obligation of a pre-triage adds friction to the system, one can only speculate the reasons for such results:

- Population was already only using the local emergency for serious occurrences (contradicting popular believe that non emergent cases are impacting performance),
- A drier/warmer period compared to previous years have lead to less emergency cases,
- The new rules lead to a change in the mix using the Manchester triage.

Just to name a few. Despite these results, it is important not to hastily conclude the ineffectiveness of such policies. As highlighted by Barros and SBE (2024b) in the article that motivated this study, there are additional considerations and questions that should be addressed before drawing any conclusions, such as:

- Is the number of emergencies the best metric to measure the impact of such an intervention?
- Do other locations that participated in the trial period show similar results?
- In the event of a reduction in emergencies, where were the patients redirected?
- Does the data reflect a like-for-like comparison, ensuring a comparable situation between different periods (e.g., same number of doctors and nurses)?
- While this analysis focuses on the mean (the first moment), shouldn't the scope of such a policy also consider the predictability of random events (the second moment)?

The list of considerations goes on. At best, this study, with its simplified approach, highlights the complexities and nuances of data and impact analysis in the context of decision-making. It serves as a reminder that first impressions can be misleading.

Appendix A: modeling the counter factual

Background

Following (Cryer and Chan 2009a) we can define an intervention as such:

> Intervention analysis, introduced by Box and Tiao (1975), provides a framework for assessing the effect of an intervention on a time series under study. It is assumed that the intervention affects the process by changing the mean function or trend of a time series. Interventions can be natural or man-made. (…)

In light of this definition a series can then be defined as such:

$$y_t = m_t + N_t$$

Where m_t is the change in the mean function and N_t is modeled as some ARIMA process, possibly seasonal (see (Cryer and Chan 2009a)) and t the moment intervention occurred. Before moment t it's assumed that $m_t = 0$. The effect of the intervention on the mean function can often captured by the use of simple functions as the **step function** or **pulse function**. On both instances a dummy variable will flag the the time the intervention takes place and we can model as $m_t = \alpha S_t$. As explained by (Cryer and Chan 2009a), *α is the unknown permanent change in the mean due to the intervention. Testing whether $\alpha = 0$ or not is similar to testing whether the population means are the same with data in the form of two independent random samples from the two populations.* In case of delayed impact of the intervention more complex modeling can be used out of the same premises.

Using intervention modeling it would then be possible to evaluate if the introduction of this new policy had any impact on the mean of the time series. The α would then provide not only the intensity of the intervention but also test is statistical significance.

Limitations

Despite being a valid approach, univariate intervention analysis using a step function does include the entirety of the information available for the analysis. Due to the nature of the series an unforeseen number of variables can impact the outcome. Some of those effects could be included in N_t and probably go beyond *ARIMA* and use other model Technics. But this would require adding a layer of complexity to the model and additional context information which is not available (eg: flu intensity and weather is expected to have an impact).

The originality of the study at end is the fact that a reference series was identified by domain experts. This means **that any number of external impact from any origin which would impact the series under study would also impact the the reference series** even if at different degrees. Not using the additional knowledge provided by a reference series would be similar to estimate a Weighted average cost of capital (WACC) disregarding the market risk.

Model using exogenous variables

Being Y_{vc} the time series under analysis, it could model as follows disregarding any reference or co-variate (see "The ARIMAX Model Muddle" 2010) :

$$Y_t = \phi_1 y_{t-1} + ... + \phi_p t_{t-p} - \theta_1 z_{t-1} - ... - \theta_q z_{t-q} + z_t$$

Above is the definition for a ARMA(p,q). Adding an additional variable or co-variate is the equivalent to do the following:

$$Y_t = \beta x_{t_n} + \phi_1 y_{t-1} + ... + \phi_p t_{t-p} - \theta_1 z_{t-1} - ... - \theta_q z_{t-q} + z_t$$

Where X stands for the co-variate or reference series (in this study this would be ULS Barcelos time series) and n for the lag. In reality a strong correlation could exist at several lag moments between Y and X therefore, the above formula could be expanded as such:

$$Y_t = \beta_0 + \beta x_{t_1} + ... + \beta x_{t_k} + \phi_1 y_{t-1} + ... + \phi_p t_{t-p} - \theta_1 e_{t-1} - ... - \theta_q e_{t-q} + e_t <=> \phi(B) y_t = \upsilon(B) x_t + \theta(B) e_t$$

where $\upsilon(B)x_t$ is the transfer function of X. This formula would them allow to expand the signals available to understand and forecast future points in Y either by using one or multiple covariates/references.

Forecasting additional point in Y requires some sort of lead/follow relation between both series. In simple terms, some form of relationship between present Y and past X is needed in order to use that signal for future prediction. But, that's not the role the reference series plays in this study. **The assumptions is that the reference series is exposed to the same external factors as the study series and therefore works as an index for external factors. In this respect what really matters is if there is a strong relationship at instant moment, meaning lag 0, since the goal is to use that information to predict an alternate reality where no intervention took place.**

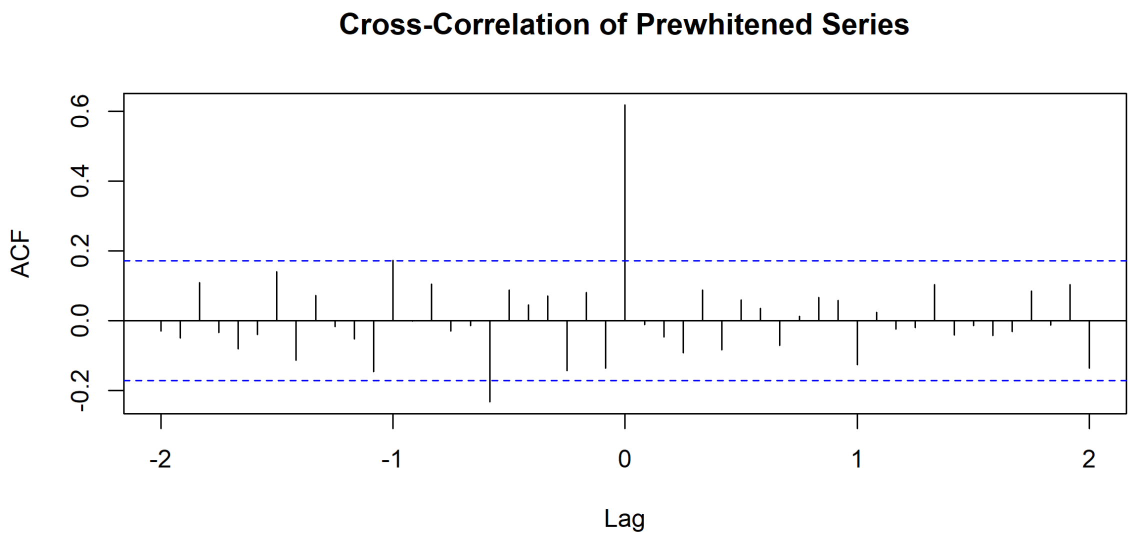


Figure 15: CCF between Barcelos and Vila do Conde

Figure 15 shows a strong cross-correlation between the series under study (ULS Vila do Conde) and the reference one (ULS Barcelos) at lag 0 just as desired[3]. This means that, until January 2024, the moment where the new policy was implemented (start of the intervention) both series followed closely each other[4].

Modeling the counter factual

Using the *auto.arima* function from package *forecast* with an external regression we can study to residuals and conclude that we have an *acceptable model* to use as counter-factual, meaning, to predict an alternative reality where no intervention took place.

[3] Both series stationary when calculations were made.

[4] The following code was used to calculate ccf. Consider vc the log variation of vila do conde and bc the log variation of bc, bot stationary series.
"‘
fit <- auto.arima(bc)
white_bc <- residuals(fit) white_vc <- residuals(Arima(vc, model = fit))
ccf_results <- ccf(white_bc, white_vc, lag.max = 24, main = "Cross-Correlation of Prewhitened Series")
"‘

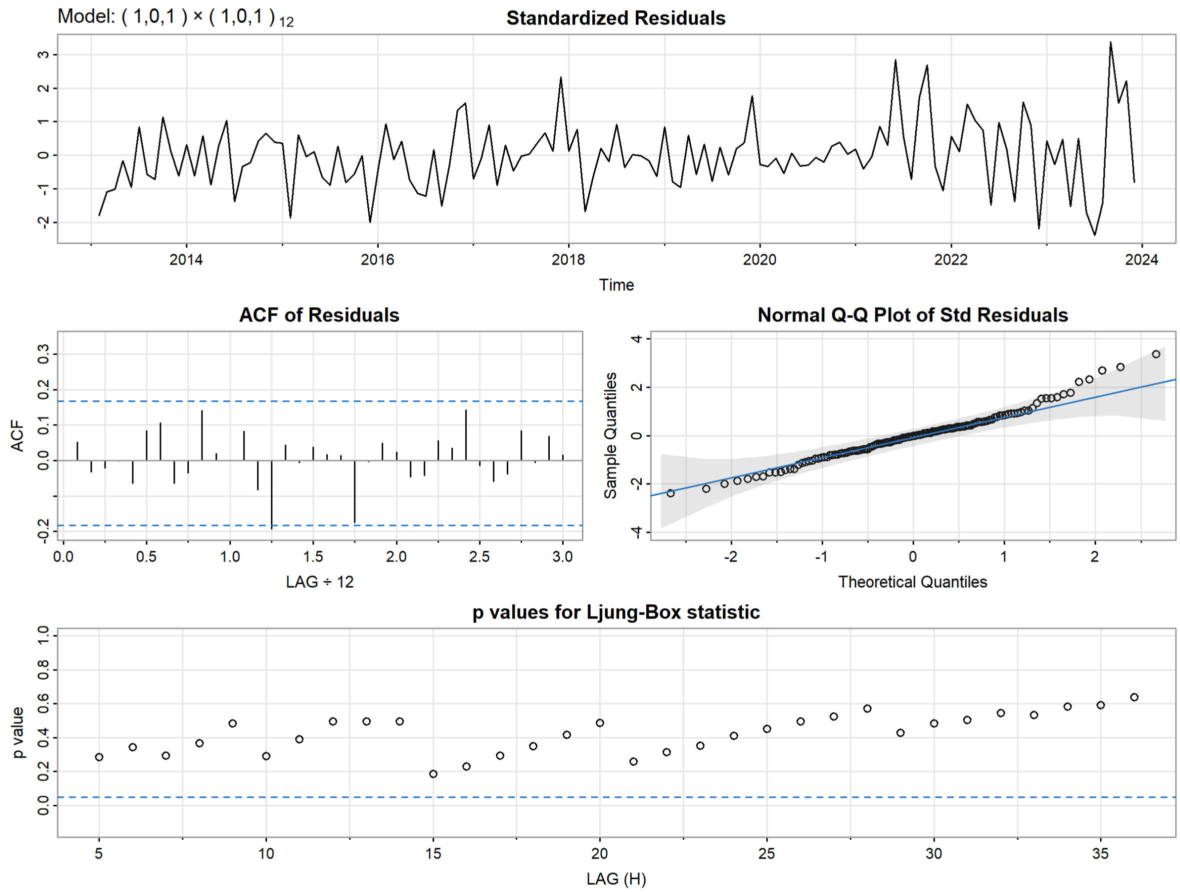


Figure 16: Counter factual model fit

Giving the way auto.arima works, by modeling to the ARMA errors, the counter factual model can be defined as follows[5]:

$$\eta = vc_t - 0.9329\, bc_t$$

$$vc_t = 0.6958\,(vc_{t-1} - \eta_{t-1}) + 0.7993\,(vc_{t-12} - \eta_{t-12}) + e_t - 0.9413\, e_{t-1} - 0.503\, e_{t-12}$$

Where VC stands for ULS Vila do Conde series and BC for ULS Barcelos series.

[5]The goal of this analysis is to use classic time series statistics, namely dynamic models to get a *first felling* on the impact of a policy. Typically, a model process would require additional care to void over-fitting, namely by splitting between test and train data. But, in this case the goal is not to predict but instead to model and understand a reality. Under this circumstances the author concluded that event of over fitting does not significantly impact the conclusions.

References

Barros, Pedro Pita, and professor na Nova SBE. 2024a. "Sobre o Programa 'Ligue Antes, Salve Vidas'." *Momentos Económicos... E Não Só*. https://momentoseconomicos.com/2024/12/18/sobre-o-programa-ligue-antes-salve-vidas/.

———. 2024b. "sobre o programa "Ligue antes, salve vidas"." https://momentoseconomicos.com/2024/12/18/sobre-o-programa-ligue-antes-salve-vidas/.

Cryer, Jonathan D., and Kung-Sik Chan. 2009b. *Time Series Analysis: With Applications in R*. 2. ed., corr. print. Springer Texts in Statistics. New York, NY: Springer.

———. 2009a. *Time series analysis: with applications in R*. 2. ed., corr. print. Springer texts in statistics. New York, NY: Springer.

Shumway, Robert H., and David S. Stoffer. 2019. *Time Series: A Data Analysis Approach Using R*. Boca Raton: CRC Press, Taylor & Francis Group.

"The ARIMAX Model Muddle." 2010. https://robjhyndman.com/hyndsight/arimax/.