PROGRAMA DE DETECCIÓ PRECOÇ DE CÀNCER DE CÒLON I RECTE

Salut/ ICO Institut Català d'Oncologia

APPLICATION OF THE ARIMA MODEL IN FORECASTING THE POSITIVITY OF A FECAL-BASED COLORECTAL CANCER SCREENING PROGRAM IN CATALONIA, 2011-2022.

Jon Aritz Panera, Noemie Travier, Ena Niño de Guzmán, Montse Garcia, Nuria Vives, Carmen Vidal, Gemma Binefa

Cancer Screening Unit,

Catalan Institute of Oncology, L'Hospitalet de Llobregat, Barcelona, Spain

COLORECTAL CANCER SCREENING PROGRAM

Northern and Southern Metropolitan Area of Barcelona

- Start year: 2000
- Change to FIT test: 2010-2011
- Extension: 2015-2016 (100,000 -> 500,000)
- Target population: 500,000

INTRODUCTION

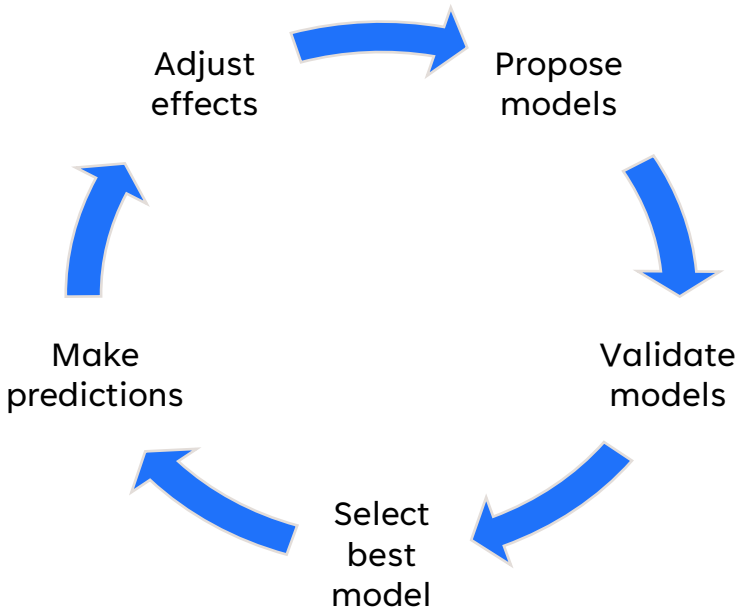
Our questions:

- Is positivity decreasing?
- Is positivity related to month-season?
- Can we predict test-positivity?

Objectives:

- Describe the evolution of the positivity on the program.
- Identify calendar effects and atypical values.
- Predict test-positivity for the next 12 months.
- Evaluate the goodness of ARIMA models to make predictions.

ARIMA MODELS METHODOLOGY



Validation tests:

- Homocedasticity
- Normality
- Independence
- Invertibility and causality
- Stability and prediction capacity

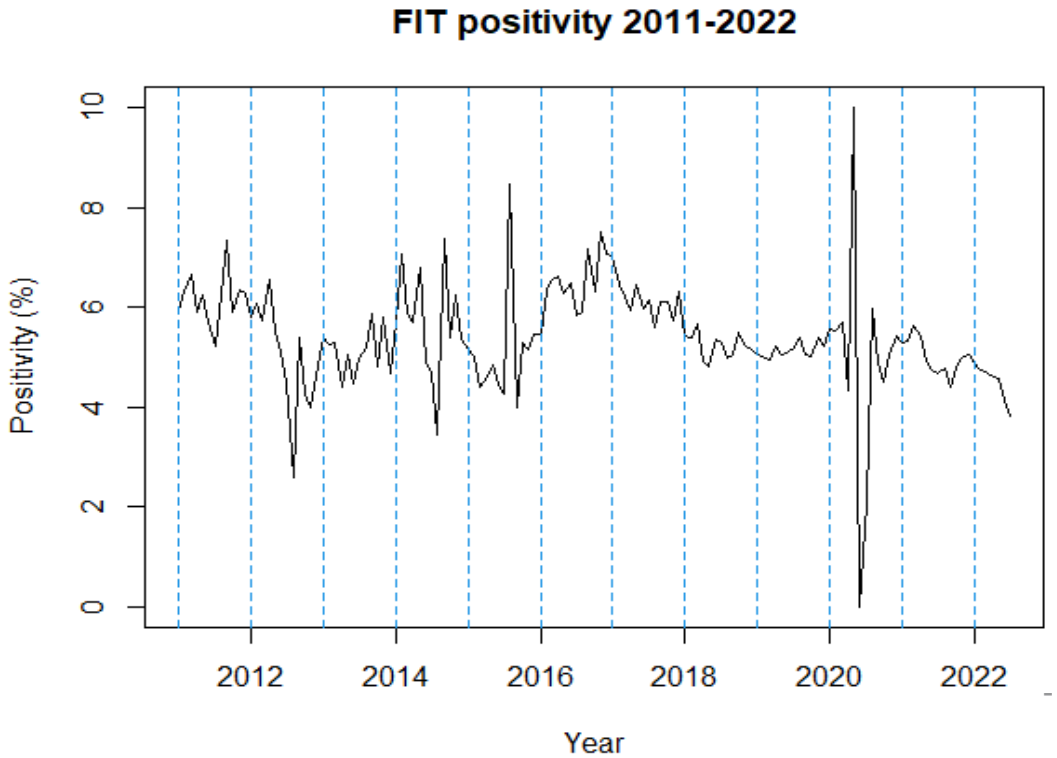
Selection parameters:

- AIC and BIC: **adequacy**
- RMSPE and MAPE: **accuracy**
- Mean length of confidence intervals: **precision**

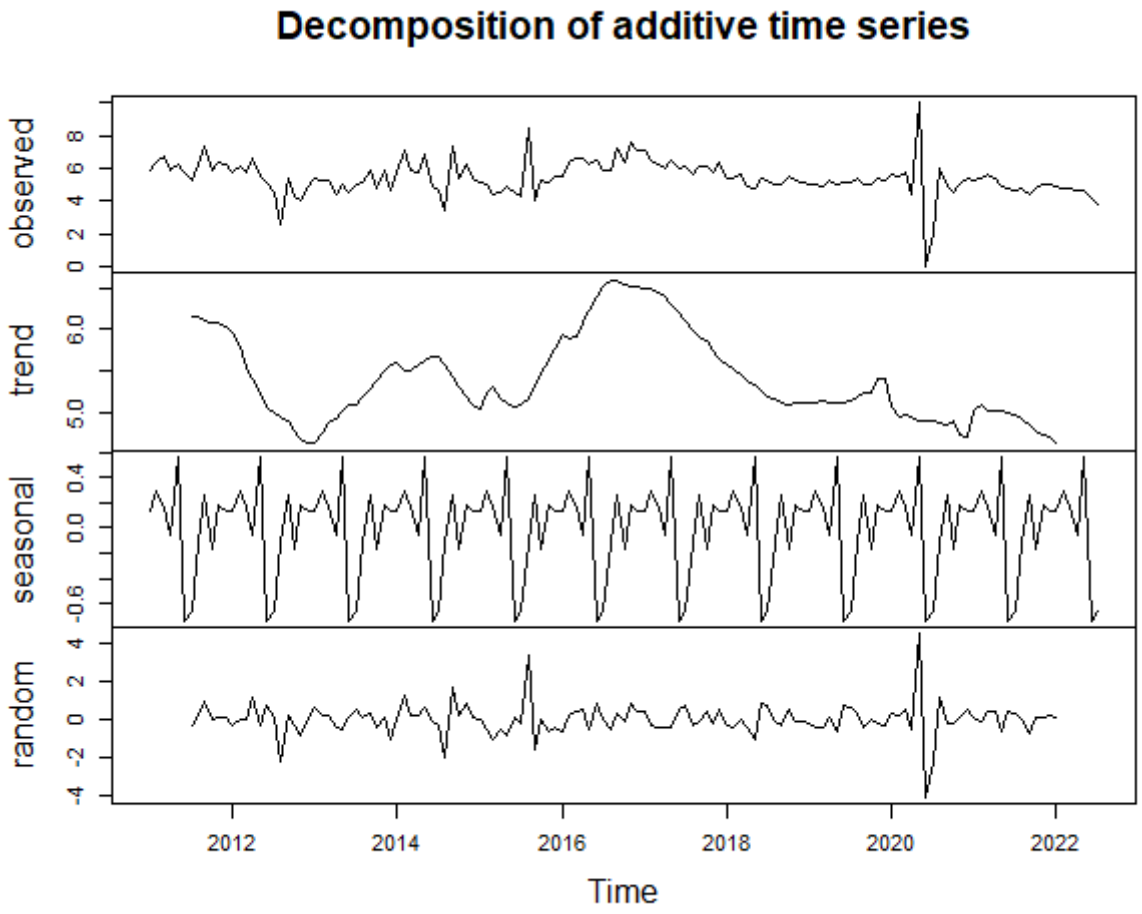
DATA AND PROCEDURE

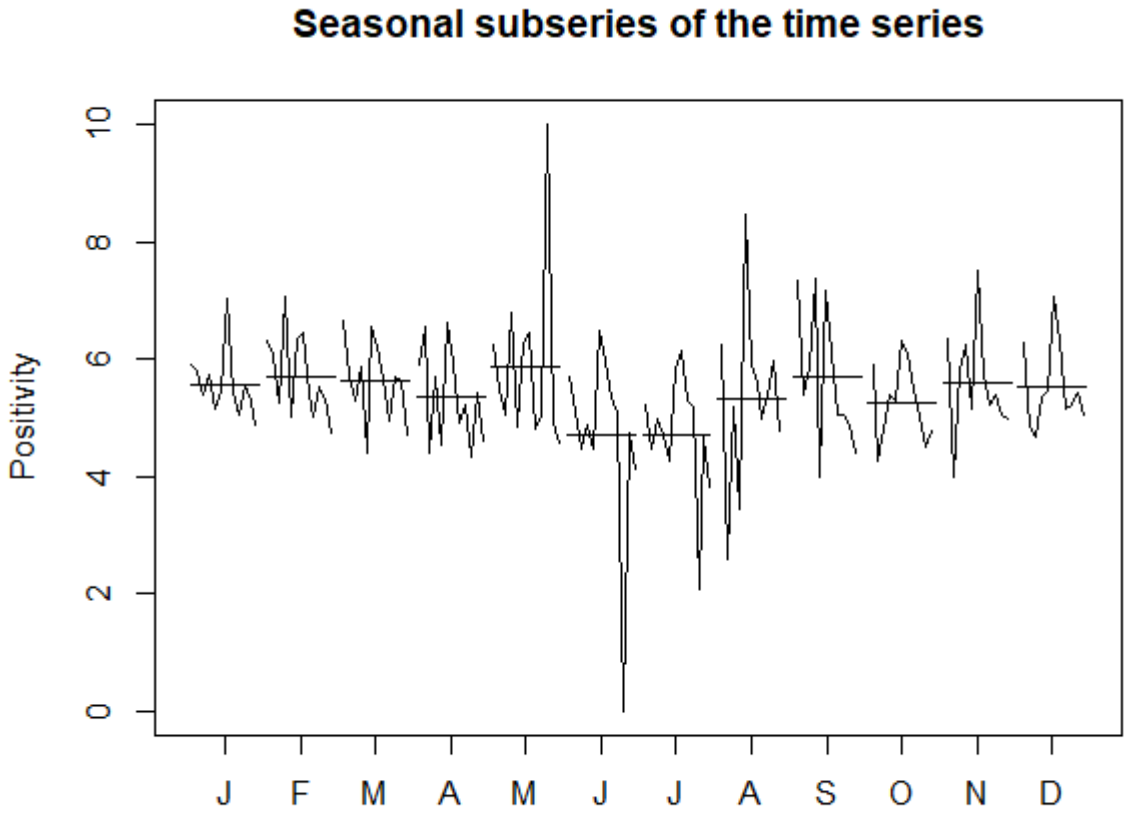
1. **Descriptive analysis**
2. **ARIMA models analysis**
   - First models
   - Estimation of 2020
   - Final models
3. **Comparison of models**

	Jan	Feb	Mar	Apr	May	Jun
2011	5.91259640	6.31016043	6.66277031	5.89743590	6.23052060	5.68278202
2012	5.78199052	6.05555556	5.74106620	6.53846154	5.48895899	5.12170385
2013	5.37992235	5.24756665	5.27497194	4.41379310	5.06329114	4.47470817
2014	5.74380165	7.06273369	5.85556278	5.69476082	6.79380215	4.87288136
2015	5.15235457	5.02306589	4.40894569	4.54908220	4.82954545	4.47042641
2016	5.44241573	6.36097561	6.54611719	6.60707346	6.28394104	6.48893931
2017	7.04225352	6.46423540	6.21326616	5.94336687	6.43325694	5.97984813
2018	5.42986425	5.40593097	5.66343042	4.91302056	4.80915348	5.34711075
2019	5.06386861	5.01679261	4.93637110	5.22166362	5.05005363	5.07999533
2020	5.55830284	5.51729597	5.71134929	4.34782609	10.00000000	0.00000001
2021	5.28890248	5.32825880	5.61858822	5.40669476	4.90309517	4.73574247
2022	4.88771466	4.73156643	4.71132691	4.59864584	4.58445491	4.10998834
	Jul	Aug	Sep	Oct	Nov	Dec
2011	5.20764667	6.25000000	7.35068913	5.88972431	6.34920635	6.25978091
2012	4.47185813	2.59740260	5.39033457	4.27112349	3.99096386	4.85519591
2013	4.99075786	5.17241379	5.85774059	4.81029810	5.80357143	4.65799413
2014	4.69361147	3.44827586	7.35294118	5.39752006	6.24082232	5.36320434
2015	4.26829268	8.45070423	4.00000000	5.28401585	5.15051331	5.47195622
2016	5.83173996	5.89080460	7.15101932	6.29385499	7.52181089	7.07489632
2017	6.14368621	5.58051549	6.09239831	6.11519919	5.73802542	6.30566945
2018	5.27589545	4.96913580	5.03727129	5.48092423	5.22388060	5.15293054
2019	5.17857143	5.39298670	5.06182827	5.02309001	5.37983002	5.21366504
2020	2.06185567	5.97284574	4.86589129	4.49614840	5.06087372	5.42642837
2021	4.68263080	4.77038797	4.38875544	4.77491730	4.98944240	5.06305289
2022	3.82369045					

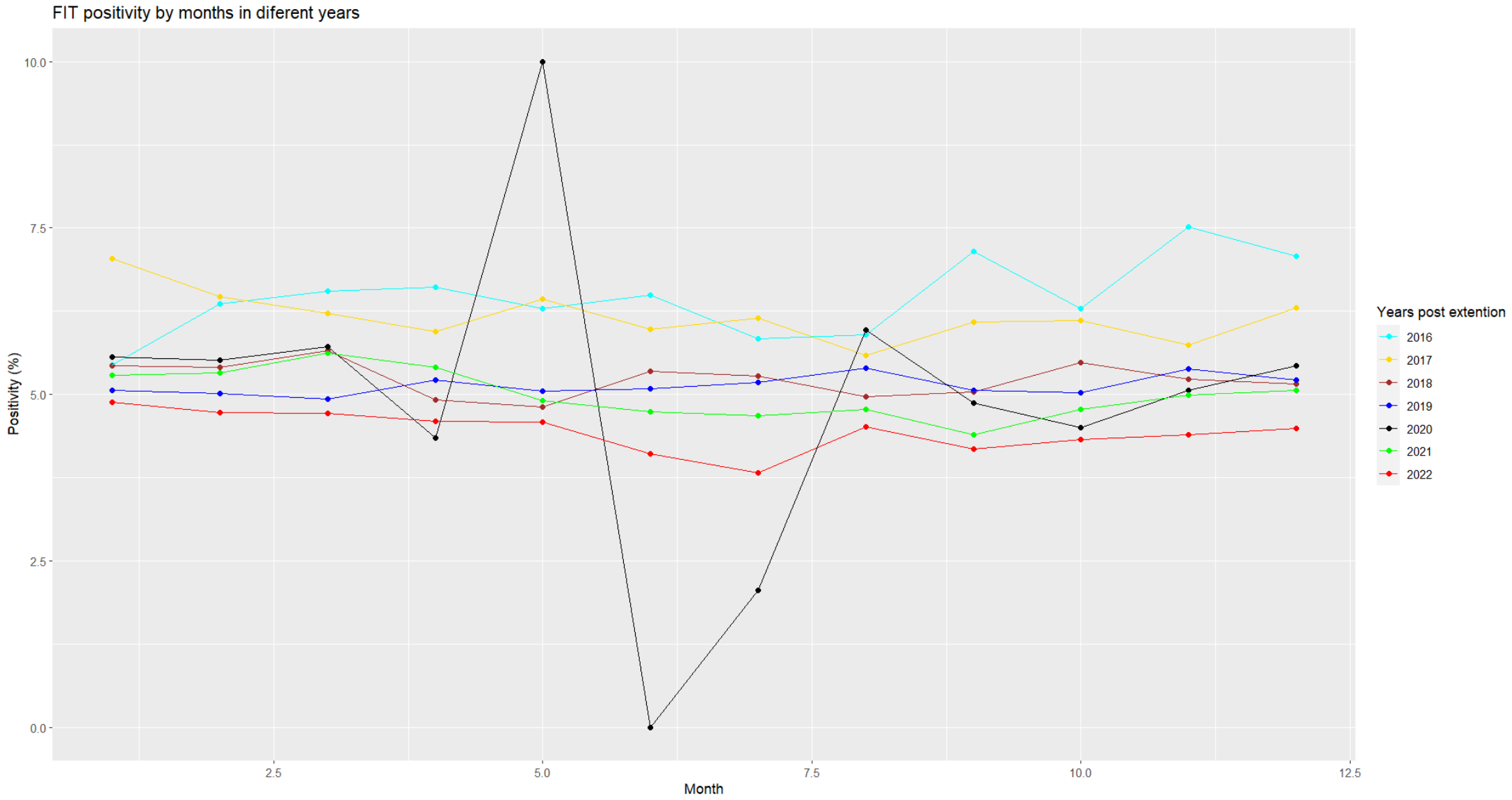


1. DESCRIPTIVE ANALYSIS



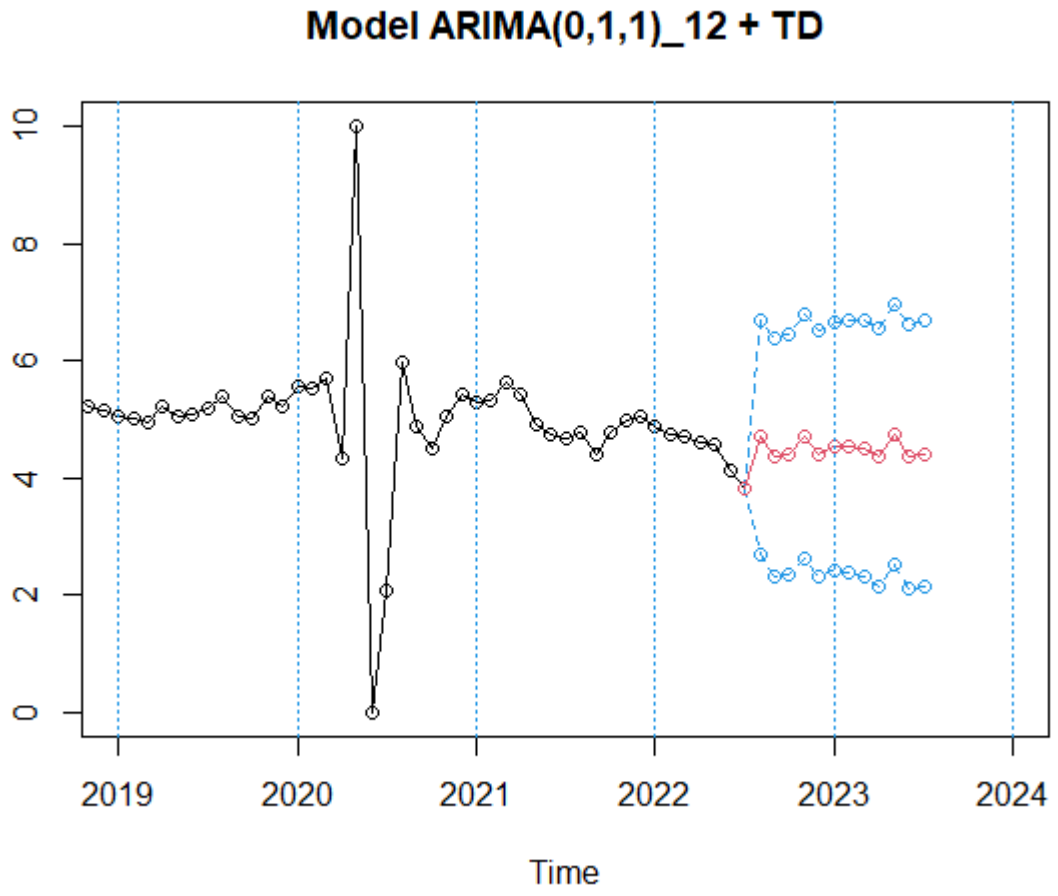


1. DESCRIPTIVE ANALYSIS



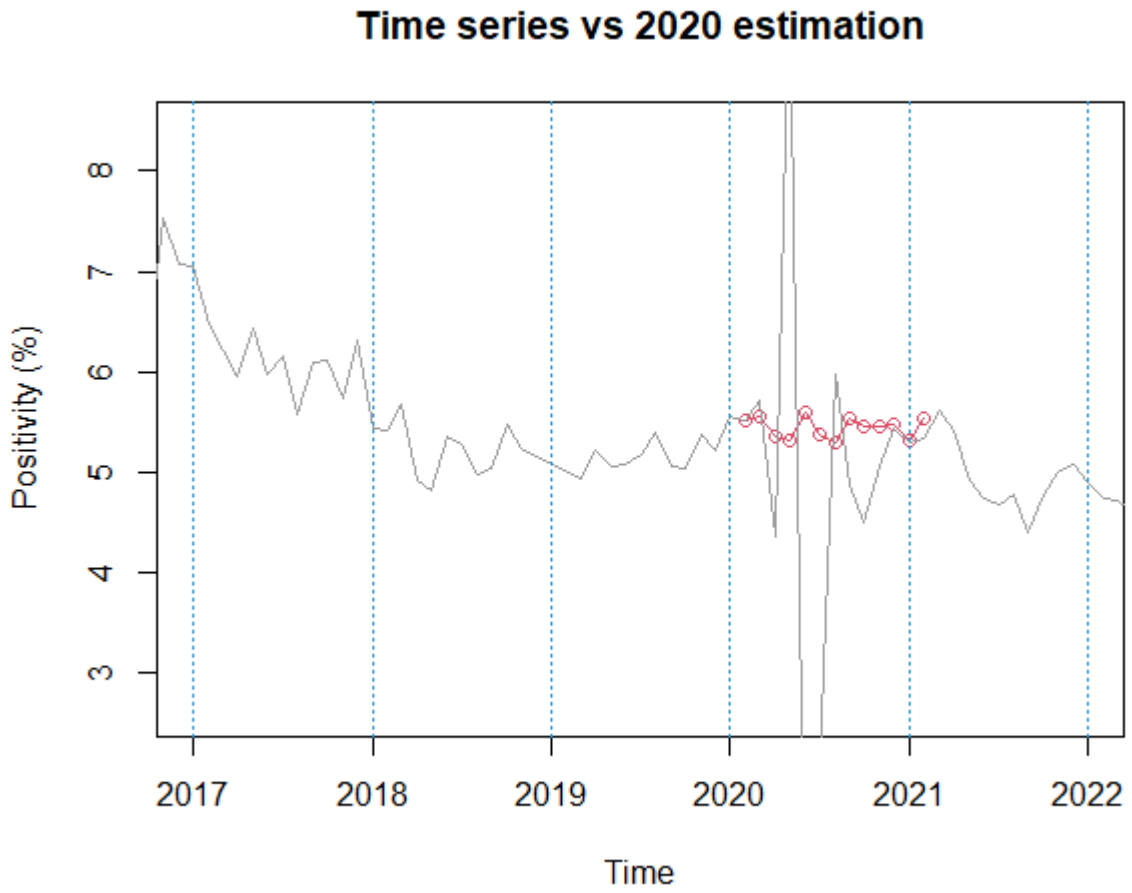
2. ARIMA model 2011-2022 (July)

Model	Coeficients	AIC	BIC	RMPSPE	MAPE	CI Mean length
ARIMA(0,1,1)_12	1	402.78	417.1	12.51 %	9 %	4.43
ARIMA(0,1,1)_12 +TD	2	401.35	410.14	9.49 %	6.76 %	4.28



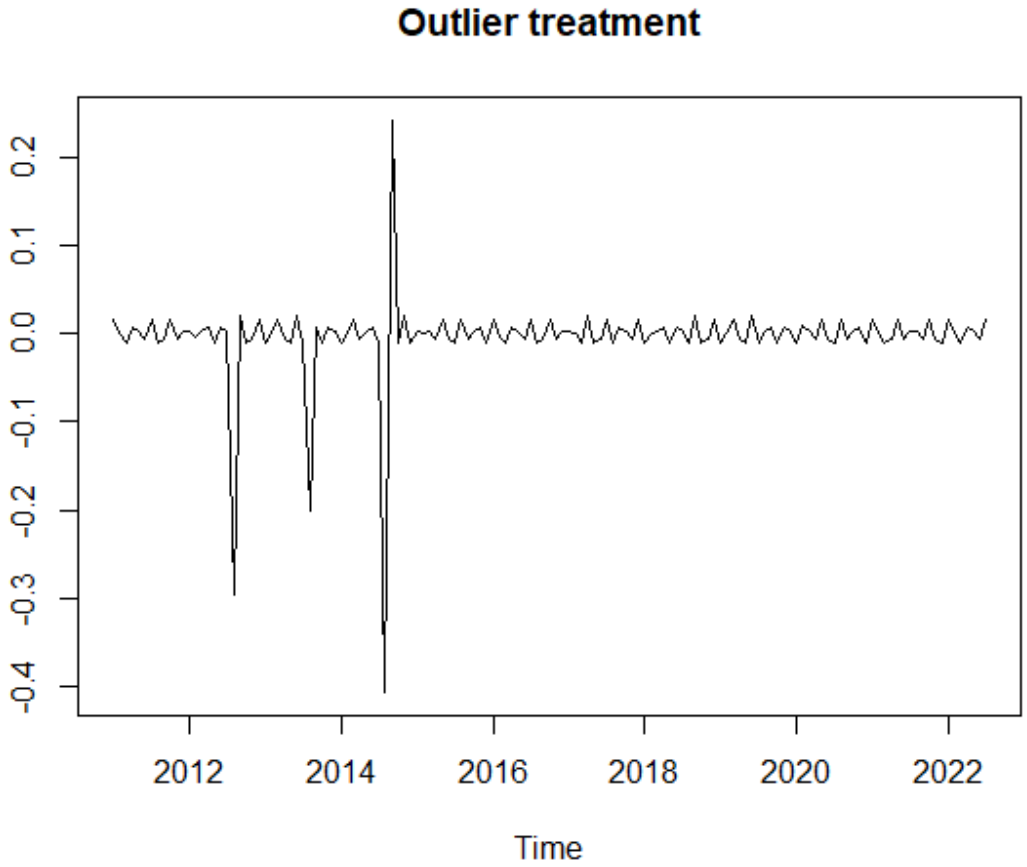
3. ARIMA model 2011-2020 (February)

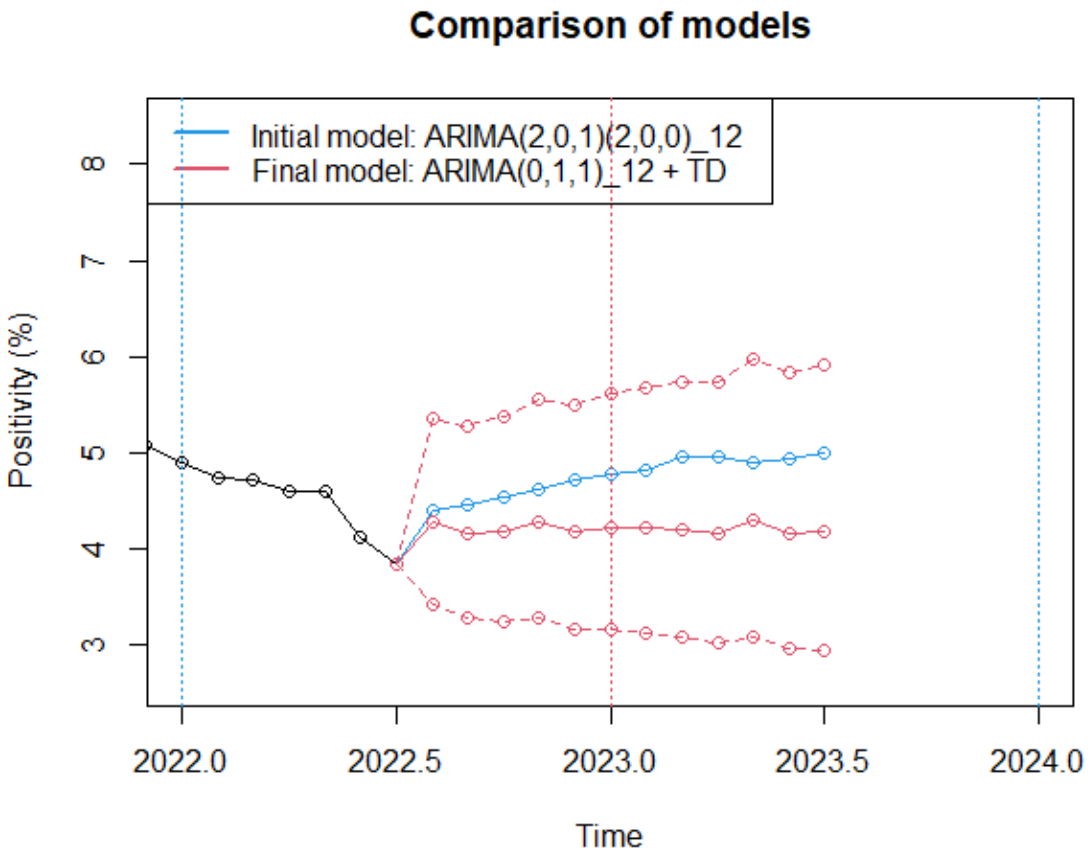
Model	Coeficients	AIC	BIC	RMPSPE	MAPE	CI Mean length
ARIMA(2,0,1)(0,0,2)_12	6	-108.89	-89.98	3.93 %	3.38 %	0.62
ARIMA(2,0,1)(0,0,2)_12 +TD	7	-140.07	-118.47	3.57 %	2.94 %	0.09



4. ARIMA model 2011-2022 (July)

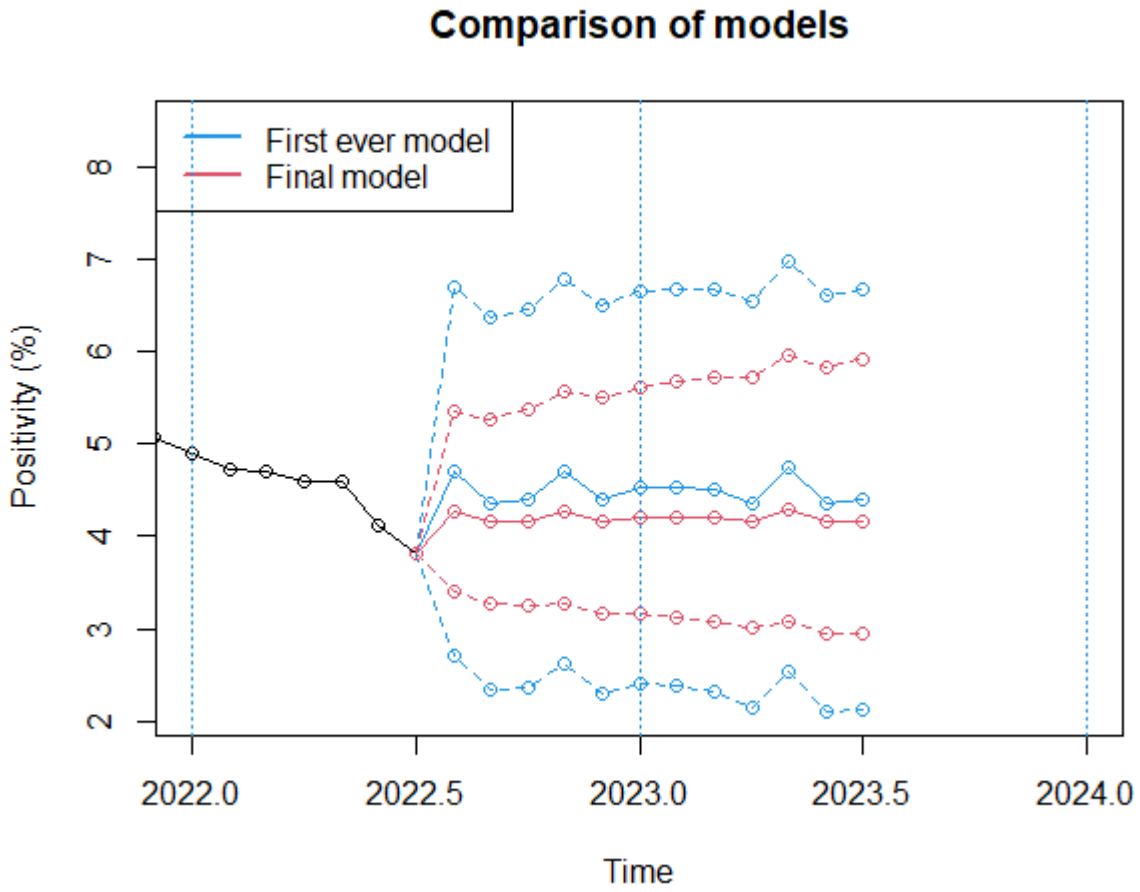
Model	Coeficients	AIC	BIC	RMPSPE	MAPE	CI Mean length
ARIMA (2,0,1)(2,0,0)_12	6	-165.08	-144.54	11.58 %	8.82 %	0.56
ARIMA (2,0,1)(2,0,0)_12 +TD	7	-199.31	-175.83	11.57 %	8.77 %	2.5
ARIMA (0,1,1)_12 +TD	2	-199.76	-190.98	7.59 %	5.34 %	0.37





5. COMPARISON OF MODELS

Month (2022-2023)	Model 1 predictions	Real data	Model 2 predictions
August	4.70 (+ 0.18)	4.52	4.27 (- 0.25)
September	4.35 (+0.17)	4.18	4.15 (**- 0.03**)
October	4.4 (+ 0.08)	4.32	4.17 (- 0.15)
November	4.70 (+ 0.31)	4.39	4.27 (- 0.31)
December	4.40 (-0.08)	4.48	4.17 (- 0.31)
January	4.53 (- 0.21)	4.74	4.21 (**- 0.53**)
February	4.53 (+ 0.36)	4.17	4.21 (+ 0.04)
March	4.50 (+ 0.35)	4.16	4.20 (+ 0.05)
April	4.35 (+ 0.28)	4.08	4.15 (+ 0.08)
May	4.75 (+ 0.04)	4.71	4.28 (- 0.43)
June	4.35	-	4.15
July	4.40	-	4.17
Over predictions	8/10	-	3/10
Under predictions	2/10	-	7/10
|Prediction|>0.1	7/10	-	6/10
Mean deviation	0.21	0	0.21



5. FINAL COMPARISON

6. DISCUSION

Conclusions:

- Time series are **good tools** to describe **evolution** of our screening programs parameters
- Even if ARIMA models are **descriptive** models they can be a **useful tool to make predictions**

Future analysis:

- Make predictions using other models
- Create a participation time series
- Create a rejection of colonoscopy time series

REFERENCES

[1] Chen, S., Wang, X., Zhao, J., Zhang, Y., & Kan, X. (2022). Application of the ARIMA Model in Forecasting the Incidence of Tuberculosis in Anhui During COVID-19 Pandemic from 2021 to 2022. *Infection and Drug Resistance*, 3503-3512.

[2] Esmaeilzadeh, N., Shakeri, M., Esmaeilzadeh, M., & Rahmanian, V. (2020). ARIMA models forecasting the SARS-COV-2 in the Islamic Republic of Iran. *Asian Pacific Journal of Tropical Medicine*, *13*(11), 521-524.

[3] Fang, L., Wang, D., & Pan, G. (2020). Analysis and estimation of COVID-19 spreading in Russia based on ARIMA model. *SN Comprehensive Clinical Medicine*, *2*, 2521-2527.

[4] Sarkar, R. R., & Chatterjee, C. (2017). Application of different time series models on epidemiological data-comparison and predictions for malaria prevalence. *SM J Biom Biostat*, *2*(4), 1022.

[5] Ibáñez-Sanz, G., Garcia, M., Milà, N., Hubbard, R. A., Vidal, C., Binefa, G., ... & Moreno, V. (2019). False-Positive Results in a Population-Based Colorectal Screening Program: Cumulative Risk from 2000 to 2017 with Biennial ScreeningCumulative False-Positive Risk in Stool-Based Screening. *Cancer Epidemiology, Biomarkers & Prevention*, *28*(11), 1909-1916.

Thank you for you attention!

Salut/ ICO Institut Català d'Oncologia

http://ico.gencat.cat