Trend analysis of total population in Spain:

A historical perspective, 1277-1838

MASTER ECONOMETRIE, STATISTIQUES

PARCOURS ECONOMETRIE APPLIQUEE

April 2024

Yava Vilar Valera

Olivier Darné

Summary

Introduction

Since the beginning of humanity, world's population has been developing and growing continuously. At a global level, more and more people inhabit the world, as well as each country or region that makes it up. From a wide range of biological to economic causes or migratory movements that shape it, the population follows a trajectory of great relevance to analyse since this trajectory has numerous implications in the structure of society, in the economy and in other areas as important such as the environment or biodiversity. The population maintains a key relationship and a similar evolution with all these mentioned factors. On the one hand, economic growth has historically driven population growth while a greater number of people stimulates demand for goods and services, employment, and economic activity in general. On the other hand, the limitation of available resources and environmental degradation seems to be greater as the population continues to grow.

Despite this global and continuous increase, there exist historical episodes in which society has experienced a considerable decrease in inhabitants because of famine, disease, or war. Likewise, there are regional differences worldwide in which the pace of population growth dissipates. Currently, the most economically advanced countries seem to increasingly limit the number of births, while in developing countries family policies are put into practice to reduce the large existent amount of reproduction.

This work focuses on understanding the dynamics of the population of Spain from the period 1277 to 1838. In general terms, it is a country that has followed a population evolution like that of the rest of the world, and more specifically, similar to European countries. However, given its economic delay in certain important historical stages relative to European countries, such as the introduction into the Industrial Revolution, the demographic increase associated with them has also arrived later. More particularly, this work will have as main objective econometric modelling whose purpose is the prediction of future values based on past values and the natural behaviour of the series. Test data will be used to compare the predicted values with the real ones, in order to evaluate the predictive capacity of the different models and to select the one that best fits the dynamics of the data and that offers a lower prediction error.

Preliminary analysis

Presentation and characterization of the series

As introduced, this work focalises on studying temporal series of total population in Spain, whose figures are expressed in millions, during the period 1277-1850. Data source comes from databases provided by “La Fundación Real del Pino”, a private non-profit organization located in Madrid, Spain[1].

Total population usually refers to the total number of inhabitants that compose a given area or territory. Different sources identify the term with the same main principle or foundation. The INSEE[2] (2023) defines it as follows: “All the persons, nationals or foreigners, established in a lasting fashion on the economic territory of the country even if those persons are temporarily absent”. The OECD (2024) characterizes it as “All nationals present in, or temporarily absent from a country, and aliens permanently settled in a country.”[3] According to the World Health Organization (2024), population refers to “All the inhabitants of a country, territory, or geographic area, total or for a given sex and/or age group, at a specific point of time”[4].

All over the studied period, demography in Spain has known a remarkable increase. Indeed, it has evolved from 4.4 million of inhabitants in 1277 up to 13.2 million in 1838, which represents a rise of 200%. In figure 1, we can appreciate a distinction in the growth speed between the dates 1277-1700 and 1700 onwards. The first period is characterized by a more gradual increase, while the second one experiences a higher growth rate. The beginning of the 1700s coincides with a period of colonial and economic expansion, together with the introduction of a new Bourbonic dynasty that brought political stability. In addition, there were improvements in the health and medicine sector that could certainly influence demography dynamics. From 1810-20, the country started

---

[1] https://frdelpino.es/investigacion/category/01_ciencias-sociales/01_economia-espanola/02_economia-espanola-perspectiva-historica/. Consulted on 31/03/2024 at 21:55

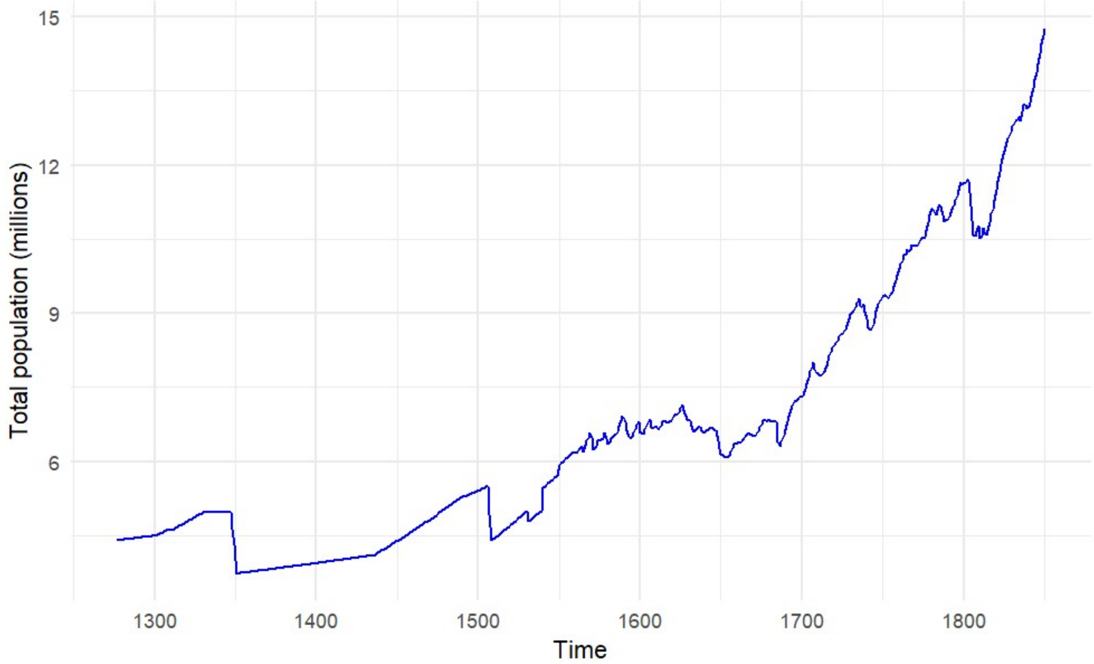
[2] (INSEE, 2023) Consulted on 31/03/2024 at 21:37

[3] https://data.oecd.org/pop/population.htm. (OECD, 2024)Consulted on 31/03/2024 at 21:40

[4] https://www.who.int/data/gho/indicator-metadata-registry/imr-details/1121. Consulted on 31/03/2024 at 21:43

to delve into the Industrial Revolution, another important event that caused a rapid demographic expansion. [5]

Figure 1 : Evolution of Total Population in Spain, 1277-1838



Source: Own elaboration with R

However, the country has also experienced some periods of demographic slowdowns. In particular, figure 1 shows a big population decrease around the years 1340-1350. This could be explained by the propagation of the Black Death that affected Spain in the year 1348, and that is estimated to have provoked a loss of between two and three million people in the country[6]. We observe a similar pattern around the beginning of the XVI century. The discovery of America in 1492 might be a factor determining this decrease, given that it generated an immigration flow from Spain to the American continent of hundreds of thousands of people, motivated, among others, by a search of economic opportunities. In the same year, the Catholic Monarchs, Fernando II of Aragon and Isabela I of Castilla, promulgated the Edict of Granada, which ordered the expulsion

[6] https://recyt.fecyt.es/index.php/IHE/article/view/86097. Consulted on 31/03/2024 at 22:27

of all Jews residing in the kingdoms of Castilla and Aragon. This measure was part of a series of events that took place during the Reconquista, the period in which the Christian kingdoms of the Iberian Peninsula recaptured territories that were under Muslim rule. The expulsion of the Jews had serious consequences for population in Spain, as many Jews emigrated to other European countries or to North Africa (Leandro Prados de la Escosura C. Á.-N.-C., 2021).

Another important population decline appears at the beginning of the XIX century. A significant event that took place in that date was the Napoleonic invasion of Spain, which triggered the Spanish War of Independence (1808-1814), responsible of a big loss of human lives due to battles, combat, and atrocities committed by the invading forces.

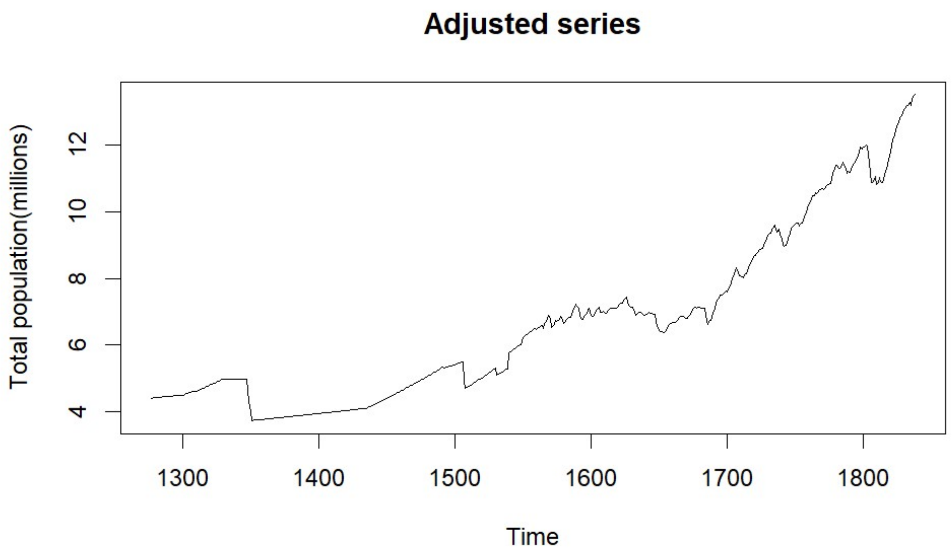
Despite different economic or social events that have upset the population trajectory, only one outlier is detected in the series. Figure 1 in annexes shows this outlier trough a red point in the series, and table 1 summarizes relevant information concerning this point. This latter represents a LS (Level Shift) outlier type in the year 1508, when population changed its trajectory level towards a rapid and natural recovery after the sharp decrease motivated by the expulsion of the Jews and the emigration flows to America just mentioned above. T. statistic being larger than 1.96, the value is significantly an outlier at the 5% level.

Table 1: Detected outlier in the series

Type	ind	Time	coefhat	tstat
LS	232	1508	-0.2967	-4.586

As outliers can distort the mean and variance, affect the quality of model fit and the reliability of predictions, among others, an adjusted series has been constructed, visible in figure 3, with which all the following analysis will be carried out.

Figure 2 :Adjusted series from outliers, 1277-1838



Source: Own elaboration with R

Series stationarity

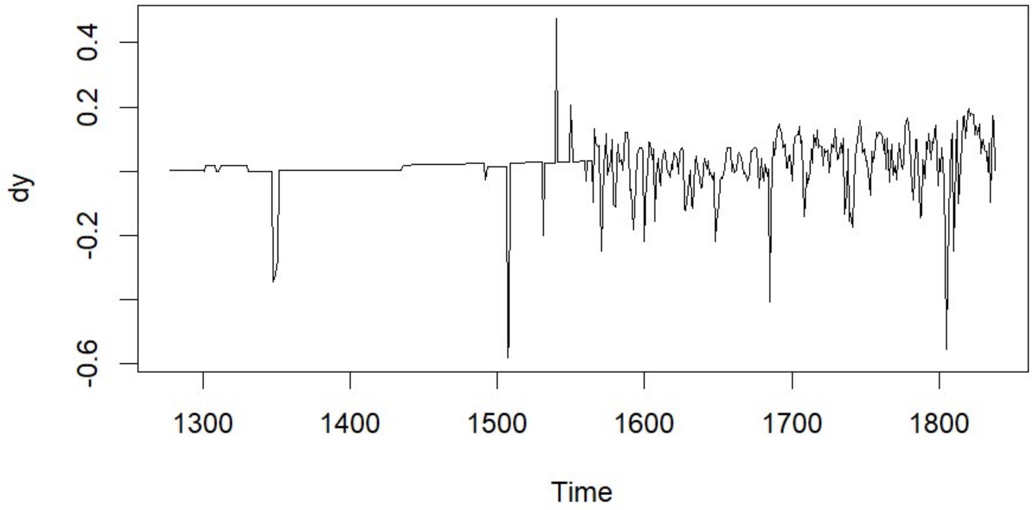
It is essential that the series is stationary given that otherwise, it cannot be modelized by some important prediction models. To verify this condition, Augmented-ADF and KPSS tests are applied. The first mentioned test hypothesis that series is not stationary. Therefore, for it to be rejected, p. value must be smaller than 0.05. Null hypothesis of KPSS test, on the other hand, stipulates series is stationary. For it to hold, the associated p.value must then be greater than 0.05. Table 2 show tests results for the raw series and after having differentiated it one time.

Table 2: Satationnarity tests

Test	Augmented-ADF test	KPSS test
	Raw series	
p.value	0.952	0.01
	Series differenciated order 1	
p. value	0.01	0.01

We can appreciate that both tests fail to consider the raw series as stationary. With a first differentiation, however, Augmented-ADF test suggests stationarity is respected, while KPSS still rejects the null hypothesis, indicating a lack of stationarity. However, with results provided by the first test, we can already take this assumption for granted.

Figure 3 : Series differentiated order 1



Source: Own elaboration with R

Figure 3 shows the graphic appearance of the adjusted series after having been differentiated one time. As we can see, now it shows a constant mean and variance, given that the values of the series tend to return to the same constant level after any fluctuation along time, and that deviations from the mean have a similar amplitude.

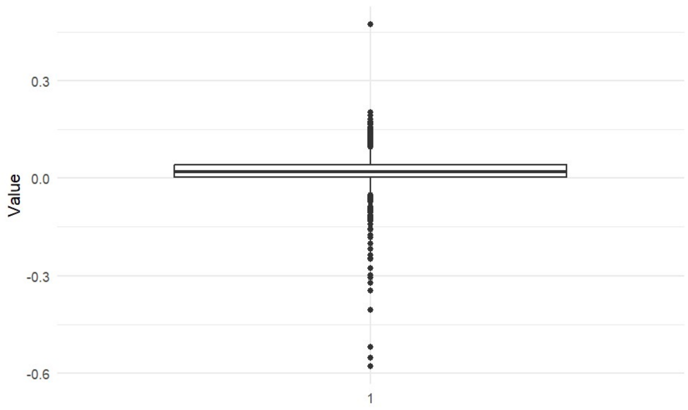
Descriptive statistics

Table 2 : Descriptive Statistics

Mean	Standard Deviation	Skenwness	Kurtosis	Shapiro.test (p. value)
0.16	0.081	-2,06	14,09	2.2e-16

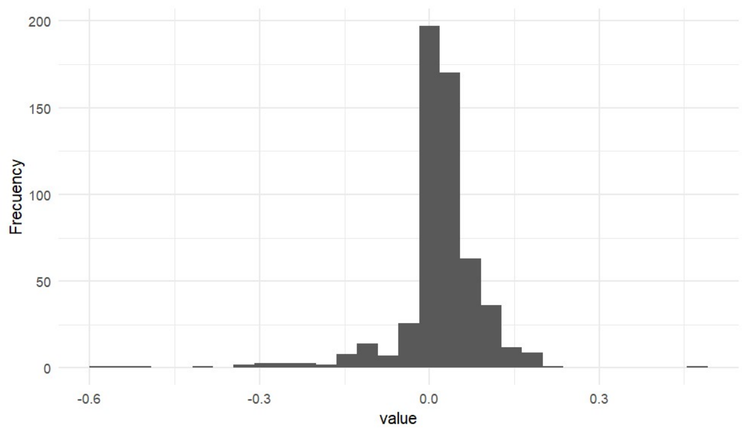
To have a better understanding of the distribution and the statistic properties of the series, table 2 provides a summary of the main descriptive statistics. We observe a mean of 0.16 and a standard deviation equal to 0.081, this latter indicating a relatively little dispersion around the mean. Skewness test indicates the series is asymmetric to the left, with a mean greater than the median, and with more values in the right than in the left of the distribution. Indeed, skewness test value is negative. Kurtosis test result being of 14,09, this indicates a pointed distribution, with a high concentration of values around the mean, and with potentially more extreme values than a normal distribution. Shapiro test, whose null hypothesis claims data follow a normal distribution, confirms what has been just described. The p.value being lower than 0.05, the series is not normally distributed. Likewise, figures 4 and 5 show the series boxplot and histogram. Their constat leads to the same conclusion: a pointed distribution with a slightly asymmetry to the left with a lot of points far of the mean.

Figure 3: Series boxplot



Source: Own elaboration with R

Figure 4: Series histogram



Source: Own elaboration with R

Linear models

In order to analyse temporal series and make future predictions, it is essential to previously estimate models that capture the dependence parameters of past values and other predictor variables. The main assumption of linear models is that the relationship between variables is linear. In addition, we have to note that AR models have been estimated with the differentiated series because stationarity is a fundamental assumption, while the remaining have been computed with the raw series because they can model a non-stationary series and automatically differentiate it.

1. **AR(1)**

To start with, AR (Autoregressive) models, with p parameters, include the values lagged from the dependent variable as explanatory variables. In this case, the current information is only explained by the immediate past value, given that p=1. If we estimate this model, AR(1), making use of different methods and software R packages, we obtain the following results:

Table 3: Estimates of AR(1) models

	OLS	YuleWalker	Burg	MLE	MLE(forecast)	CSS(tseries)
AR(1)	0.4829	0.4829	0.4829	0.4921	0.4820	0.4829
σ 2	0.005063	0.005073	0.005054	0.005054	0.005054	-

We can observe that the coefficients remain almost identical for all the different methods used, varying from 0.4820 to 0.4921. A parameter lower than 1 indicates a stationary series that tends to revert its mean, that is, that tends to converge towards a mean value in the long term. A positive coefficient reveals an increase of each current value as long as time increases. Indeed, a positive term AR(1) implies that when the value of the time series t-1 is greater than its mean, the value at time t is also greater than its mean, suggesting a pattern of growth over time.

On the other hand, sigma squared represents the variance of the residual errors, which are the differences between the observed values and the values predicted by the model. A low σ^2 suggests that the residual errors are small and that the AR model fits the data well, while a high one indicates the opposite. We can see on table 3 that the variability of observations non-explained by the different models is quite low, around 0.00506, suggesting good quality models. If we had to choose one among them following σ^2 criteria, that would be either Burg or MLE methods together with MLE from forecast package.

2. AR(p)

It is also possible to automatically detect the optimal number of parameters in an Autoregressive model. To do this, different information criteria such as AIC, BIC or Hyndman and Khandakar algorithm are often considered in R. If our interest focuses on minimizing AIC information, an automatically detected AR model with again one parameter has been reached, whose coefficient with the variance of residuals errors are shown on table 4.

Table 4: Estimates of AR(p) model

AR(1)	σ 2
0.4829	0.005073

The interpretation of the new parameter is the same as before because it has the same value than much of the previous methods employed to detect AR(1) coefficients, and an identical sign (positive). Consequently, it seems that Autoregressive model that best fit the data is one with only one parameter, which indicates only a positive short-term relationship with its immediately previous value, and no one with value two, three or more periods ago. The reduced number of parameters can be a reason that supports this fact, considering that parsimony criterion is better satisfied. The variance of the residual errors is small, 0.005037, slightly lower than that of the previous AR(1) models.

3. ARIMA(p,q,d)

We can also estimate complete ARIMA models with the auto.arima function of forecast package using R software. This function takes into account AIC and BIC information criteria to select the best suitable model. After having applied this function to the raw series, we obtain an optimal model with a differentiation rang of order 2, ARIMA(0,2,2). That is, the series would have to be differentiated two times.

Table 5: Estimates of ARIMA(p,d,q) model

	MA(1)	MA(2)
coefficient	-0.5602	-0.4167
t-stat	-16.08460	-11.95592

The model is characterised by having two parameters Moving Average processus significant at the 1% level (t-stat greater than 2.64). This shows a dependency relationship between the value at a specific point of time t and error term both at the immediate previous period, t-1, and at the second previous period t-2. Negative coefficients indicates this relationship is negative and the greater coefficient of MA(1) compared to MA(2) suggests that the error term at t-1 has a higher effect on the current value.

4. Holt-Winters

Appeared in the 1960s, Holt-Winters method is a prognostic technique that can capture both tendance and seasonality in data, with the advantage of being able to adapt to no-seasonal series. Updates to the level, trend, and seasonality components are based on three smoothing parameters: α, β, and γ, where α (alpha) controls the smoothing of the level component, β (beta) controls the trend component, and γ (gamma) the seasonality component. As the series analysed on this work does not have a seasonal component, Holt-Winters model has been estimated in R specifying an inexistent gamma parameter.

Table 6: Estimates of Holt-Winters model

α	1	a	13.5114
β	0.4401	b	0.0602

In the instances when α and β parameters are close to 1, we talk about a "smooth" forecast, while if they are close to 0, they reflect a "strong smoothing". In the first case, recent past has a big weight on forecasting. In the second case, a higher importance is given to distant past. In his case, we could say near past has a bigger weight than remote past, because α is completely equal to 1 and β coefficient is relatively important. The model coefficients are:
"a" is the estimate of the level in the last period, which seems to to be 13.51137000, "b" is the estimate of the slope of the trend in the last period. In this case, 0.06019208. These coefficients can be interpreted as the last observation's estimates of the level and slope of the trend, respectively. In summary, this model seems to be indicating that the current level of the series is approximately 13.51, and the current trend is increasing by approximately 0.0602 units per period.

5. ADAM-ETS

The ADAM-ETS method is an approach that combines exponential smoothing (ETS) approaches with state space models to model time series with error, trend, and seasonality components, also allowing the inclusion of other covariates.

Table 7: Estimates of ADAM-ETS model

	Estimate
alpha	1.0000 *
beta	0.999 *
phi	1*
level	4.3716 *
trend	0.0179*
drift	1.0039*
other	0.2509 *

Table 7 exhibits parameters associated with the calculated model. Firstly, we can observe only significant parameters. The smoothing coefficient alpha controls the relative importance of past observations in estimating the current level in the ETS model. In this case, the estimated value of alpha is 1.0, specifying that full weight is given to the most recent observation in the calculation of the current level. Beta captures the relevance of differences between past observations in estimating the current trend in the ETS model, which has a value of 0.9999. This indicates that differences between past observations have a very important weight in approximating the current trend. Phi represents the smoothing coefficient associated with the trend. A value equal to 1 suggests a complete weight to recent observations when estimating the trend, which means that the trend will adjust more quickly to recent changes in the data. The model also has a significant level component equal to 4.3716. The latter represents the average level of the series at a given time, it is an assay of the baseline around which the data fluctuates. Trend and drift components have a coefficient of 0.0179 and 1.0039 respectively. The first component models the general direction in which the data changes over time, which is upward because of a positive sign, but not very pronounced. The second component models gradual or systematic changes that cannot be fully captured by the trend component. The value of 1.01 indicates that the time series is increasing by 1.0039% per year. Finally, the significant presence of another unknown parameter or component in the model is specified with a value of 0.25.

6. ADAM-ETS – ARIMA

By combining the ETS and ARIMA models in the ADAM-ETS-ARIMA approach, we aim to capture both exponential smoothing patterns and autocorrelation patterns in the data.

This allows modelling a wide range of time series structures such as trends, seasonality, abrupt changes, and other complex patterns.

However, it has turned out that this approach has produced a model with identical parameters and with the same coefficients as the simple ADAM-ETS model, which may suggest that the latter is adequate to capture the structure of the data and that the inclusion of additional ARIMA components does not provide a significant improvement on predictive capacity. This seems coherent with the fact that ADAM-ETS has projected an ARIMA(0,2,0) model, that is, with 0 AR or MA processus.

7. CES

The time series CES models proposed by Svetunkov, Kourentzes and Ord in 2022 refer to "Complex Exponential Smoothing" models. They are an extension of traditional exponential lysing methods that are based on the theory of complex variables and offer an innovative way to model and forecast time series. They can predict both stationary and non-stationary series, and can be adapted to seasonal data through their formulation in a state space form. Auto.ces function of R software estimated a CES(n) model, where *n* represents the non-seasonality term.

8. SSARIMA

SSARIMA models are a prediction technique that combines the State Space with ARIMA methodology for the analysis and forecasting of temporal data, whose acronym meaning is: "State Space ARIMA". In this way, these models can include ARIMA components to model the time structure of the series, as well as additional state space components to shape other unobservable factors that can influence the time series. Auto.sarima function in R preconised an ARIMA(0,0,3) model with a constant. The coefficient values are specified on table 8.

Table 8: Estimates of SSARIMA model

Intercept	MA(1)	MA(2)	MA(3)
0.0161	0.5009	0.2134	0.0755

The ARIMA(0,0,3) structure means that there is no autoregressive (AR) component but three moving average (MA) terms. Moving average coefficients represent the contribution of the residual values in the three previous periods to the prediction of the current value. We can observe that the weight of each previous period decreases with the higher number of parameters. That is, the influence of the most distant periods is smaller compared to the closest ones. Indeed, coefficient of MA(1) is equal to 0,5 while that of MA(3) equals 0,075.

All the parameters and coefficients associated with each method is summarised on table 9. AIC (Akaike Information Criterion) and AICc (Corrected Information Criterion) have been included in order to compare models and be able to determine which one seems to be better. The AIC seeks to minimize the discrepancy between the model and the data, penalizing the complexity of the model. This means that a good model will be one that has a good fit but with a minimum number of parameters.

The AICc is a correction to the AIC, especially useful when the sample size is small compared to the number of parameters.

Table 9: Models parameters summary

	AR(1)	AR(2)	ARIMA (0,2,2)	LED Holt-Winters	ADAM ETS	SSARIMA	CES
Ar1	0.4820	0.4921	0.3904				
Ma1			0.1440			0.5009	
Ma2						0.2134	
Ma3						0.0755	
Alpha				1	1		
Beta				0.4401	0.9999		
Phi					1		
Distribution	Normal	Normal	Normal		Normal Generalised	Normal	Partial
AIC	–1367.96		–1344.38		-3395.009	-1367.81	–1231.551
AICc			–1344.33		-3394.748	-1367.70	–1231.508

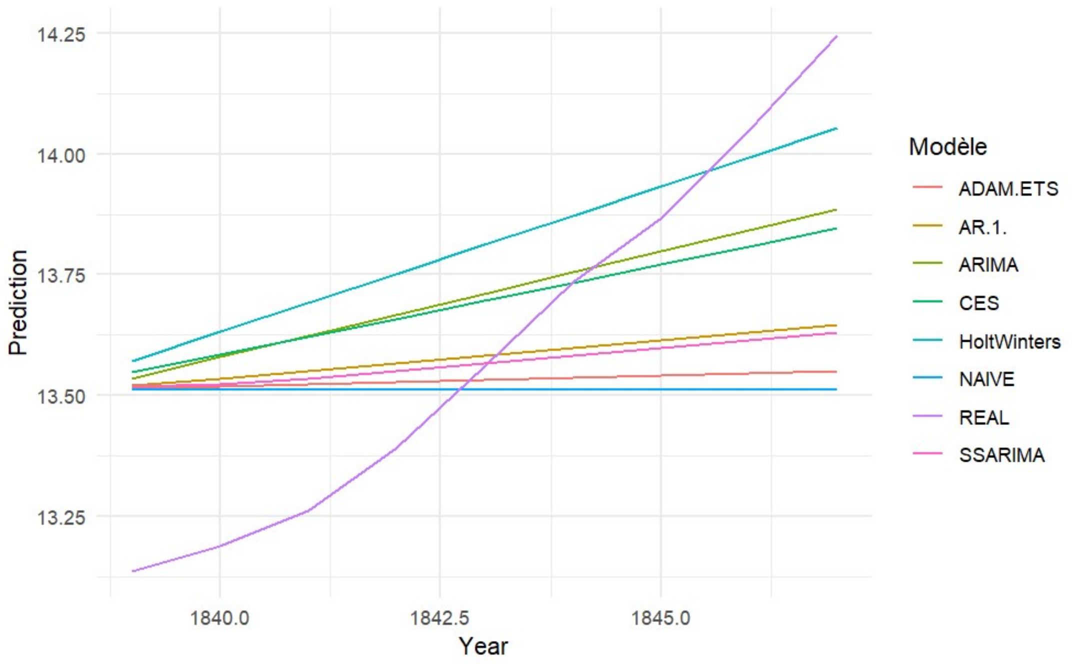
ADAM-ETS offers the lowest AIC and AICc. This is the reason why this model is preferred to others if we consider this criterion. However, differences between different methos are small. The only one that slightly moves away is CES, with the highest AIC and AICc, but the difference with respect to the other techniques is still small.

Linear predictions

In this section, we will use the statistical models we have previously fitted to make predictions for the nine years following the end of the time series. These predictions will be based on the patterns and trends identified in the historical data, as well as the statistical relationships captured by the models. We will explore the predictions generated and evaluate their performance based on appropriate evaluation metrics. For this latter, in addition to the computation of MSE (Mean Squared Errors), we will use a naive prediction as a benchmark to calculate R^2OOS and DM (Diebold-Mariano) test, explained later. A naive prediction is a simple forecasting approach based on the assumption that the future value of a time series will be equal to the last observed value in the series. In other words, this method assumes there is no change in the time series and that the future value will be equal to the last observed one. Thus, this approach does not incorporate any underlying trends, seasonality, or patterns in the data, and is useful as a reference to determine whether more complex models provide any significant prediction improvement.

Every forecast generated with each different method is visually summarised on figure 5. We have previously remarked that AR(1) and AR(p) models led to the same result (because automatically detected AR(p) suggested only one parameter with the same coefficient and characteristics). Thus, predictions were also the same and have been summarized into one single AR(1) model. The same happened with ADAM-ETS and ADAM-ETS-ARIMA, the reason why only ADAM-ETS model appears in figure 5. Graphics concerning forecast from each method are available in annexes, together with the exact values provided. Naive predictions and real observations have been included with the aim of comparing results.

Figure 5 : Models comparison



Source: Own elaboration with R

We can appreciate that any model has reached to exactly fit the same evolution with the same growth rate as the real values. In particular, true series increases at a considerable higher speed. In addition, all the predictions start with a higher value than the actual one in the year just after the last observation of the series. In other words, population has experienced a decrease in the year 1839 with respect to 1838, but all the models have forecasted a rise from which the series continues to grow. This slowdown could be unexpected, because it suddenly appears after a big increase without stop since the last of the War of Independence. This fact may explain why models have failed to predict a decrease. Then, following the downfall, the series could grow at a faster rate than expected to return to normal levels from before.

In any case, the models that seem to better fit the real series are Holt-Winters, ARIMA and CES models, because they are the ones that have predicted a higher increase from 1838 to 1847. We observe the preference of ARIMA, which only contained MA components, compared to models with only an AR component. This suggests that the time series is more influenced by past errors than by past values. That is, the data may exhibit erratic or random behaviour, where future values depend primarily on past errors and not on a clear trend or pattern. Likewise, we can observe that the Holt-Winters and CES methods have managed to fit the data relatively well. In effect, they are designed to capture the trend and seasonality of the series (that in this case, there is no seasonality). The series grows rapidly over time, with a clear increasing trend, and these models, in particular Holt-Winters, are the ones, with ARIMA, that have the most predicted a pronounced increase with a higher growth rate. It appears that ADAM-ETS, on the contrary, even if it had the lowest AIC and AICc, presents higher error rates. As it is not completely evident to correct the series from outliers because it only presented one level shift, and that models adapt to adjusted values, which in the future (the no-analysed data) would differ from real ones, possibly affecting the accuracy of the forecast, a complementary analysis has been carried out in annexes, figure 8. However, results show that prediction is better with the series corrected from outliers.

In order to complement the graphic assessment, we are going to consider MSE and $R^2 OOS$ performance evaluation metrics, to continue with Diebold-Mariano test. In the context of forecasting, we want our model predictions to be as close as possible to the actual values observed in the data. The MSE is a statistical measure that provides a measure of how well the model achieves this goal. Indeed, this metric offers a quantitative assessment of the accuracy of a prediction model relative to the actual values. It is based on the difference between model predictions and true observations, squared to penalize larger errors, and then the average of these squared differences is calculated. It is mathematically expressed as follows:

$$\mathrm{MSE} = \frac{1}{H} \sum_{h=1}^{H} (y_{t+h} - \widehat{y}_{t+h})^2 = \frac{1}{H} \sum_{h=1}^{H} e_{t+h}^2$$

Where H represents the forecast horizon, y_{t+h} the observed series, $\hat{Y}_{t+h}$ the prediction of a model and e_{t+h} is the prediction error.

The out-of-sample R-squared, on the other hand, is a measure of the predictive ability of a model when applied to new data that were not used to train the model. In essence, it measures how much variability in the dependent variable the model can explain in data not used in the fitting process. A higher out-of-sample R-squared value indicates a better predictive ability of the model on new data. It is also a valuable tool for comparing and evaluating the predictive performance of different models relative to a reference or benchmark model. If the value of this metric is greater than 0, it will be interpreted that the model in question has a superior predictive ability to the benchmark model. If it is less than zero, it will be the opposite, while if it is equal to 0, there will be no predictive difference. It is calculated as indicated on the following formula:

$$R^2_{OOS} = 1 - \frac{MSE_i}{MSE_0}$$

Where MSE_i represents the MSE of a model I and MSE_0 that of the reference model.

The DM test, or Diebold-Mariano test, is a statistical test used to compare the predictive accuracy of two time series models. This test was proposed by Francis X. Diebold and Robert S. Mariano in 1995, whose purpose is to determine whether there is a significant difference in the predictive performance between two models. The null hypothesis is that both models have the same predictive accuracy, while the alternative hypothesis is that at least one of the models offers a more accurate forecast. In the context of this work, the DM test will be used to compare each of the models studied in this work with the naive forecast, which helps to determine if the model in question is more interesting than making a simple prediction.

Table 10: Evaluation indicators summary

Model	MSE	R^2OOS	DM statistic	DM p.value
AR (1)	0.1109604	0.2472992	0	1
ARIMA	0.07964322	0.4597396	0	1
Holt-Winters	0.09185754	0.3768838	0	1
ADAM-ETS	0.1359391	0.07785635	0	1
CES	0.08574753	0.418331	0	1
SSARIMA	0.1130467	0.2331472	0	1
Naive	0.1474164	0		

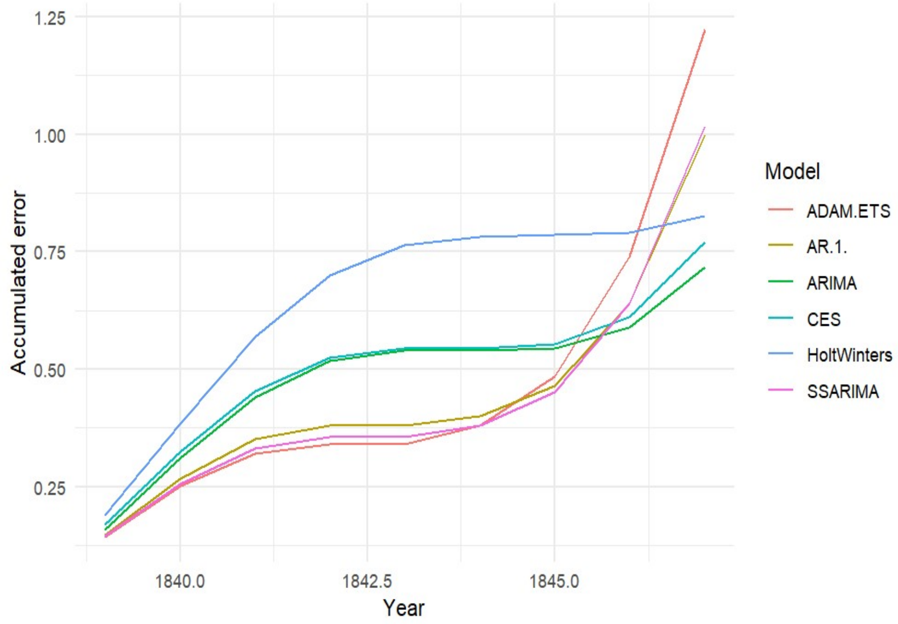
If we look at table 10, we can judge that all the models seem to provide better prediction results compared to a simple naive forecast. Indeed, the R^2OOS is always positive and even some models have relatively elevated values such as ARIMA or CES, between 0.4 and 0.5. Equally important, the Mean Squared Error of each model is lower than 0.147 , the MSE of the naive model. However, if we look at the DM test statistic and p. value, we can see that null hypothesis is accepted, meaning that there is no significant difference in the forecast accuracy between model i and the naive prediction. In any case, if we had to choose a better model on which we could comparatively rely, it would be ARIMA, because it presents the highest R^2OOS value and the lowest MSE.

Finally, another forecast performance evaluation technique is the CSPE (Cumulative Squared Prediction Error), often used to evaluate the accuracy of predictions in time series models. It is calculated by adding the squared errors of the predictions over a given period. Squared errors are used to give more weight to large errors, reflecting the importance of avoiding large deviations in predictions. Mathematically, it is expressed as follows:

$$\mathrm{CSPE} = \sum_{h=1}^{H} (y_{t+h} - \widehat{y}_{t+h})^2 = \sum_{h=1}^{H} e_{t+h}^2$$

Following this formula, whose acronym meaning has already been explicated, accumulated errors of predictions of each model have been computed, reflected in figure 6.

Figure 6 : Cumulative squared prediction errors (CSPE)



Source: Own elaboration with R

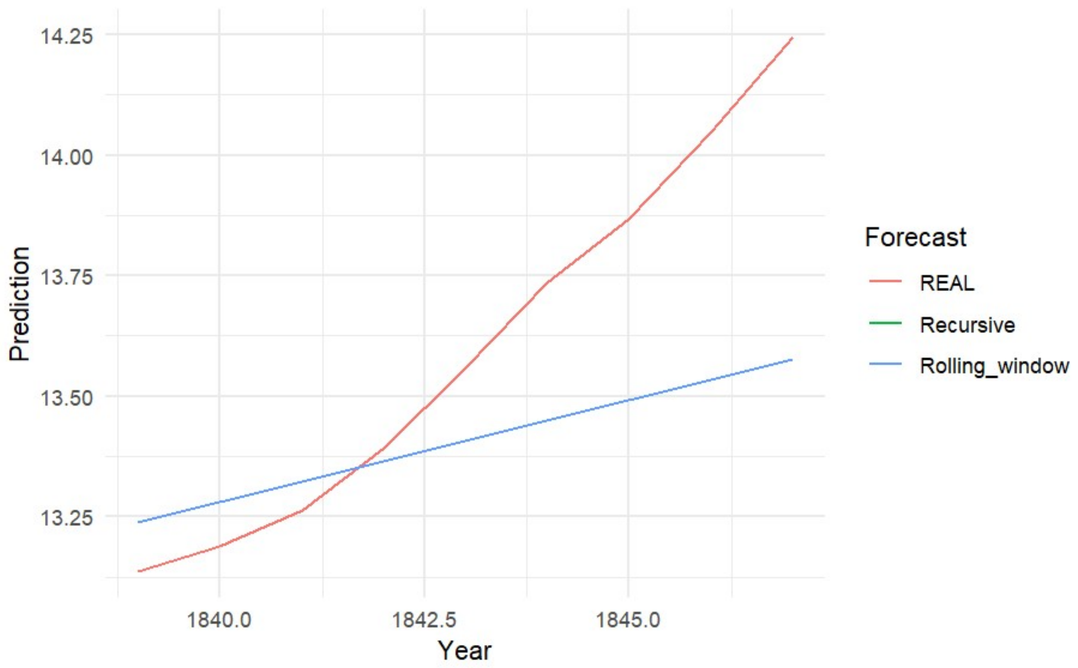
We can interpret that, again, ARIMA seems to be the best model because the cumulated sum of errors is the smallest one among the totality of CSPEs, even if at the firsts stages, it starts with a higher prediction error than other methods such as SSARIMA, ADAM-ETS and AR1. On the contrary, these last models finish having much highest CSPE values. This constant can suggest that there exists a difference on the prediction ability between models in function of the forecast horizon. At the very short-term, ADAM-ETS, SSARIMA and AR1 provide better results, whereas from a horizon of about 7 or 8 periods, ARIMA, CES and Holt-Winters become more accurate.

Predictions with ARIMA

In time series analysis, there are two main approaches to make predictions: recursive and rolling window methods.

Recursive methods use previous predictions as input to predict the next step in the time series. This means that each new prediction is based on previous predictions. However, these methods can be susceptible to error propagation and bias accumulation over time. On the other hand, the rolling window method, at each step, the prediction slides forward, removing the oldest observation and adding a new one. This process is repeated iteratively to make continuous predictions over time. We are going to make these predictions with the ARIMA model, because it has turned out to be the most suitable one. Results are shown on figure 7, and exact values are available on table 2 in annexes.

Figure 7: Recursive and rolling window predictions ARIMA model



Source: Own elaboration with R

We appreciate that the two forecast methods provide almost the same values, reason why only a blue curve appears on figure 7 masking the line of the other prediction. This can be explained by the fact that the model employed to make forecast is the same. Likewise, we can see that there still exists an important deviation from actual values and the prediction. As a positive point, these techniques have predicted a decrease in t+1 with respect to the last observation t, something that the other models did not achieve to forecast.

Conclusion

In conclusion, this work analysed the trend of the evolution of the Spanish population in history, as well as some of the most important events that caused a decline in population. We have chosen and selected different prediction models to evaluate which one is better adjusted to the structure of the series and which can be more reliable for use in a real-life situation. The result is that the ARIMA model with 3 Moving Average processes has a minor prediction error. Therefore, the work culminated in the use of two common forecasting techniques with this better model. All prediction techniques predict a tendency of increase for the near historical years, adjusting to the real life, with a satisfactory estimate of the error term. Nevertheless, the analysis and exploration of other techniques might continue to get models with predictions closest to the real values.

Discussion

As a discussion, it is important to note that the series is constructed from equations, except for the period 1750 onwards. This is observed by the mechanical behaviour of the series, and its construction process is specified in this reference (Leandro Prados de la Escosura A.-N. a.-C., 2022). Although the collected data reflects the outputs of the model rather than direct observations, equations may not fully capture the complexity of the phenomenon they seek to represent. Therefore, careful attention is required when interpreting the results of this analysis. It is relevant to consider how the modelling process may bias the conclusions, particularly when extrapolating the results to historical events or making future predictions. Additionally, cross-validation with other independent data sources is recommended to confirm the robustness of the conclusions drawn from this constructed series.

Bibliography

INSEE. (2023). Récupéré sur https://www.insee.fr/en/metadonnees/definition/c1751.

Leandro Prados de la Escosura, A.-N. a.-C. (2022). Growth recurring in preindustrial Spain Cliometrica. Appendix.

Leandro Prados de la Escosura, C. Á.-N.-C. (2021). Growth recurring in preindustrial Spain? *Cliometrica*.

OECD. (2024). Récupéré sur https://data.oecd.org/pop/population.htm

World Health Organization. (2024). Récupéré sur https://www.who.int/data/gho/indicator-metadata-registry/imr-details/1121.%20Consulted%20on%2031/03/2024

Data source:

https://frdelpino.es/investigacion/category/01_ciencias-sociales/01_economia-espanola/02_economia-espanola-perspectiva-historica/

Annexes

Figure 1 : Outliers in the series

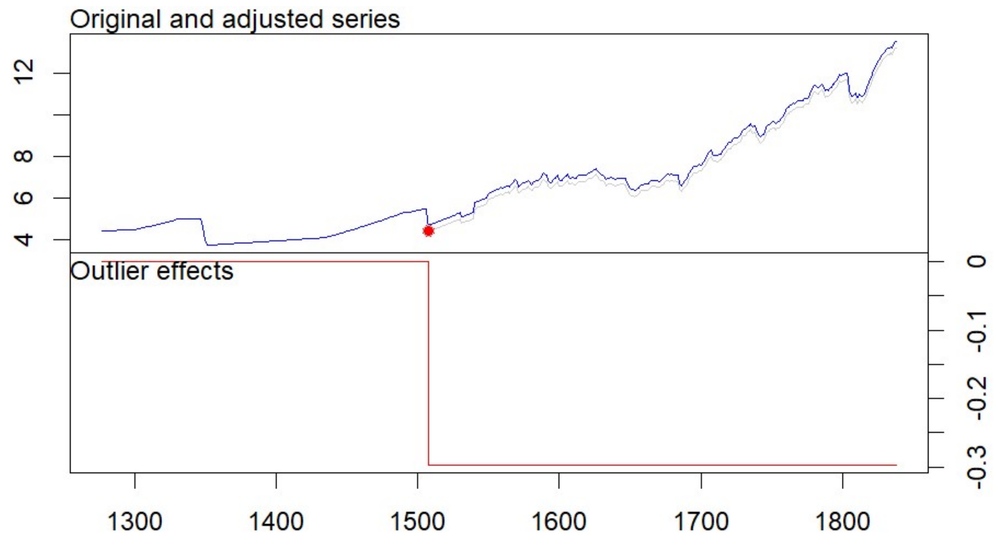


Figure 2 : Forecasts from AR(1)

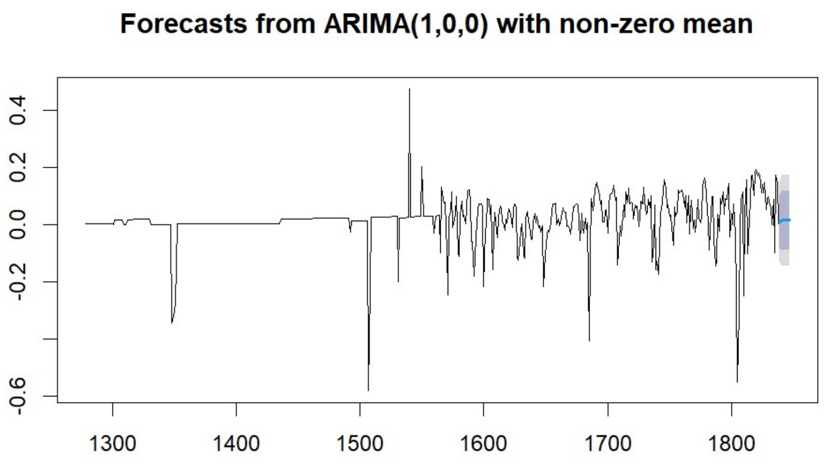


Figure 3 : Forecasts from ARIMA

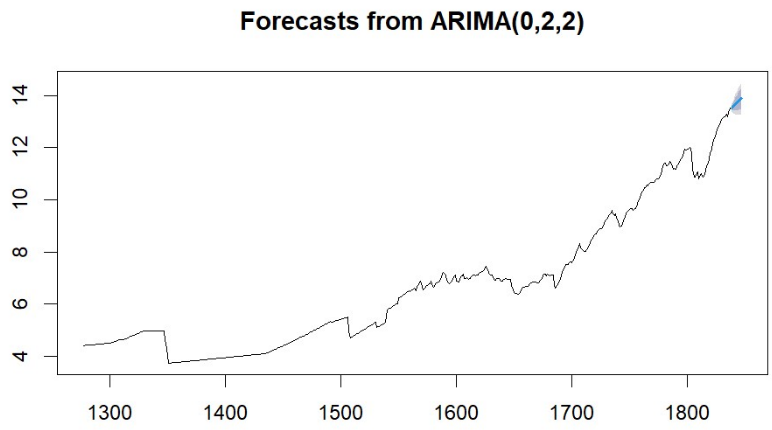


Figure 4 : Forecasts from Holt-Winters

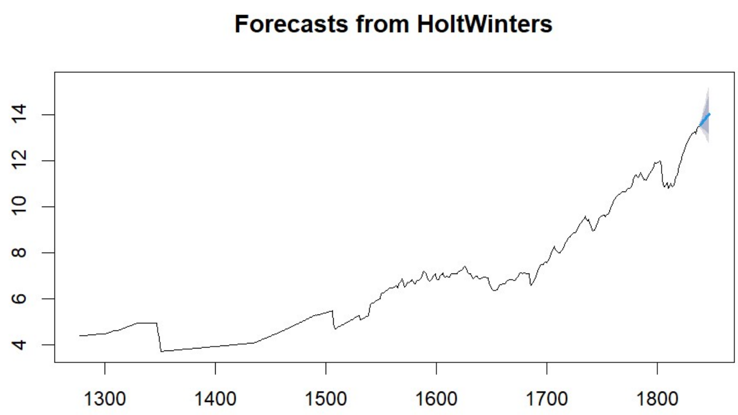


Figure 5 : Forecasts from ADAM-ETS

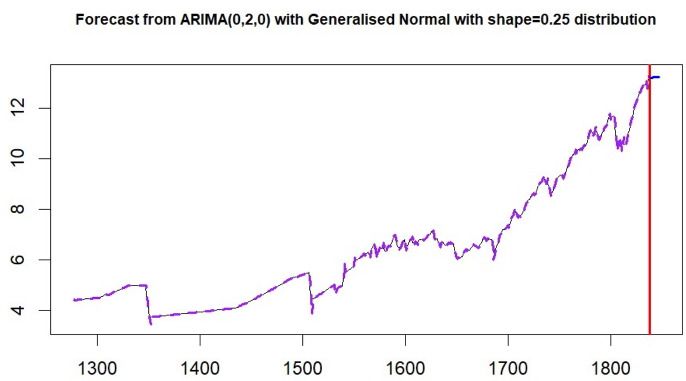


Figure 6: Forecasts from CES

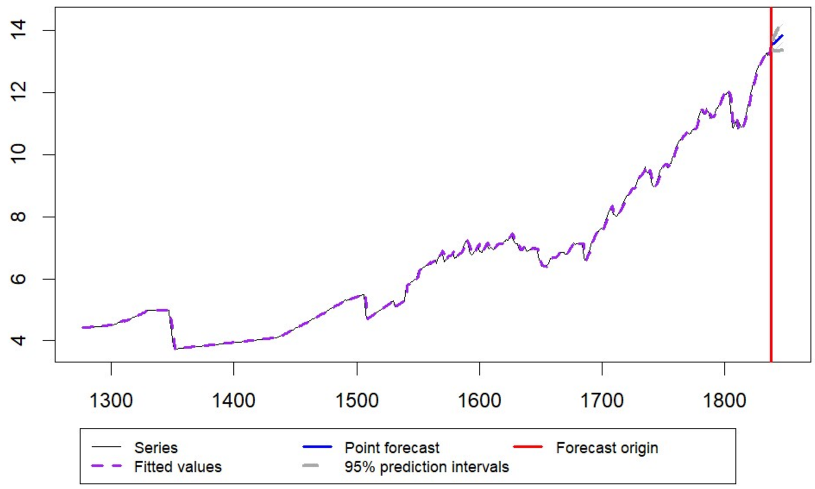


Figure 7: Forecasts from SSARIMA

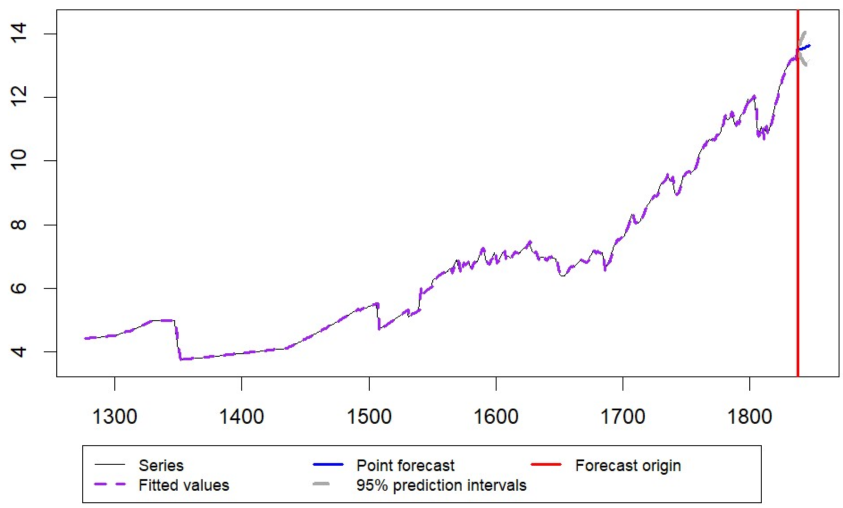


Figure 8: Forecasts without correcting the series from outliers

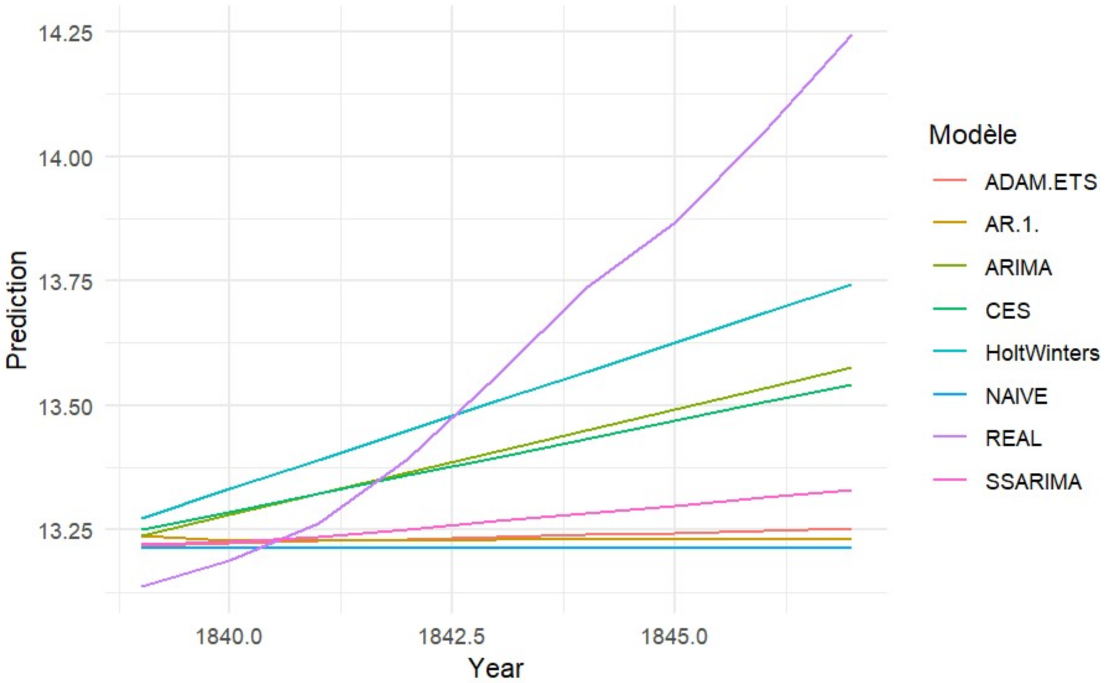


Table 1: Forecasts from every model and actual values

YEAR	AR.1.	ARIMA	HoltWinters	ADAM.ETS	CES	SSARIMA	NAIVE	REAL
1839	13.52183	13.53552	13.57156	13.51563	13.54794	13.51871	13.51137	13.13696
1840	13.53520	13.57938	13.63175	13.51989	13.58467	13.52340	13.51137	13.18958
1841	13.55017	13.62324	13.69195	13.52415	13.62174	13.53424	13.51137	13.26224
1842	13.56575	13.66710	13.75214	13.52841	13.65868	13.55038	13.51137	13.39167
1843	13.58167	13.71095	13.81233	13.53267	13.69594	13.56652	13.51137	13.55883
1844	13.59775	13.75481	13.87252	13.53693	13.73308	13.58265	13.51137	13.73500
1845	13.61391	13.79867	13.93271	13.54119	13.77055	13.59879	13.51137	13.86631
1846	13.63010	13.84253	13.99291	13.54545	13.80789	13.61493	13.51137	14.04963
1847	13.64631	13.88639	14.05310	13.54971	13.84557	13.63107	13.51137	14.24530

Table 2: Recursive and rolling window forecasts

Year	Recursive	Rolling_window
1839	13.23807	13.23807
1840	13.28034	13.28034
1841	13.32260	13.32262
1842	13.36487	13.36490
1843	13.40712	13.40718
1844	13.44937	13.44946
1845	13.49162	13.49174
1846	13.53386	13.53402
1847	13.57610	13.57631

Contents