First assignment Switzerland

Leonardo Melo 4690923 / 1940988
FMDA
Country: Switzerland
3rd November 2017

Observed data set

The observed dataset for this assignment has been obtained through the historical Statistics of the World Economy in Bolt (2014)[1]. The assignment concerns the population of Switzerland from years 1 to 2009 A.D.

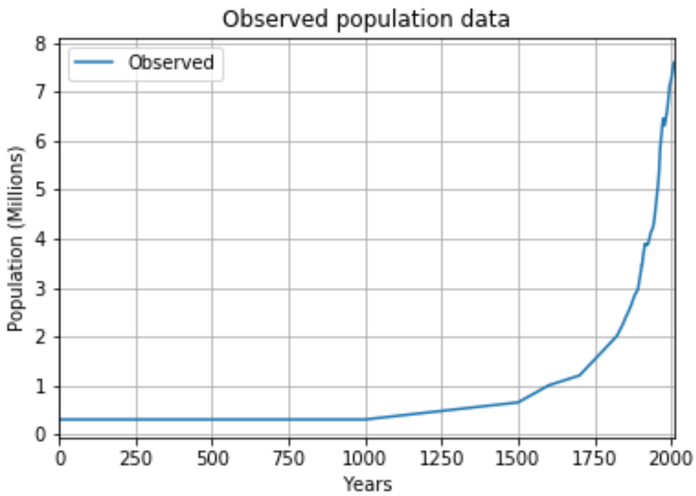

Figure 1: Observed dataset from 1 to 2009 A.D.

The growth rate indicates the pace at which a dataset evolves between two different periods. Equation (1) shows the used growth rate formula:

$$GR = \frac{P(t_2) - P(t_1)}{P(t_1) \times (t_2 - t_1)} \times 100 \ [\%] \tag{1}$$

Applying equation (1) into the observed population of Switzerland from 1 to 2009 A.D. yields the following growth rates:

Table 1: Growth rates and population in observed data

Period (Years)	Growth rate (%)	Population (millions)
1-1700	0.177	0.3 – 1.2
1700-1900	0.875	1.2 – 3.3
1900-1950	0.845	3.3 – 4.69
1950-2000	1.096	4.69 – 7.27
2000-2009	0.515	7.27 – 7.60

[1] http://www.ggdc.net/maddison/maddison-project/home.htm

Hubbert's growth model

In this model, the accumulated population $Q(t)$ depends on the value of constant K, growth rate r, peak time t_{peak} and time t. As described in Scherer (2017a), these constants influence the shape of the "S" shaped logistic growth curve. Hubbert's formula is shown in equation (2):

$$Q(t) = \frac{K}{1 + e^{r(t_{peak}-t)}} \tag{2}$$

K defines the maximum population for the growth model. The growth rate r is the pace at which the population grows until a certain plateau. t_{peak} is the year at which growth attained its peak and from that date onwards the growth rate starts to decline to attain the plateau level of K.

Calibration procedure

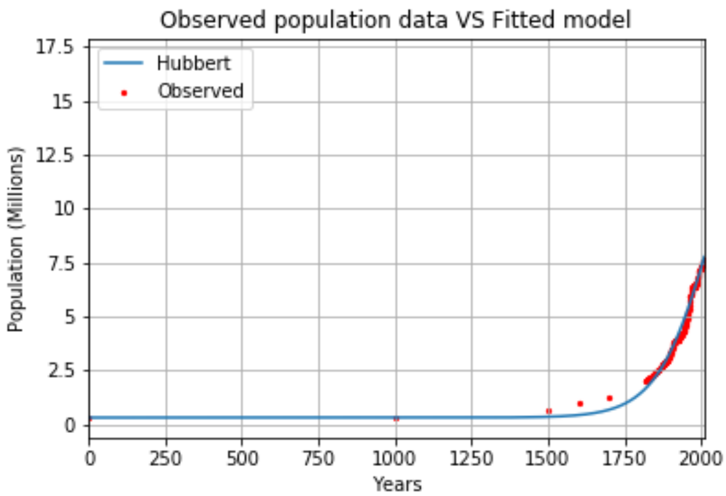

Figure 2: Observed data plotted with Hubbert's model for the year range 1-2009 A.D.

The calibration procedure took into account the growth model of Hubbert (equation 1). The observed data from 1-2009 A.D. for Switzerland showed a very steep growth in the years after 1950 with a growth rate greater than 1 while in other periods it was 20% smaller. The visual calibration was accompanied by constant robustness check with the given model evaluation tools. For every parameter change, R and the percentage bias are calculated. The calibration procedure is:

1. Fix the maximum population K to obtain a similar positive steep slope with the observed data in the period 1950-2009. The selected value is close to 17 million inhabitants.
2. Select the peak year t_{peak} to move the "S" shaped curved towards the observed data. The value 2030 was chosen to by visual inspection.
3. The coefficient r requires extensive manipulation of decimal cases to reach a satisfying fitting. The value 0.01151871 was chosen.
4. Given the fact that the population remains fairly low in the early years, a coefficient of 0.3 million was added in the Hubbert's formula.

The final equation is shown in (3):

$$Q(t) = 0.3\times10^6 \frac{17\times10^6}{1 + e^{0.01151871(2030-t)}} \tag{3}$$

Discussion

The implemented Hubbert's growth model was done using visual fitting. Given the multitude of parameters to tune, it would have been better to use automated fitting techniques allowing faster goal seeking.

Initially, the calibration started by using the Verhulst model as described in Bacäer (2011). This model also uses parameters K and r and these can be estimated automatically from 3 equally distant data points. As mentioned in the Bacäer (2011), the prediction of K using this method only applies for values separated by a few decades. In the case of the population of Switzerland, K was being estimated as close to 7.5 million. With Hubbert this largely increased to almost reach 17 million.

Equation 3 yields rather high fitting results with a $R^2 = 0.9851$ which is very close to perfect relationship as mentioned in Scherer (2017b). The percentage bias of equation 2 is 1.082% which comes close to the perfect prediction standard of 0% percentage bias.

One important aspect of the observed dataset is the dispersion of measurements across years: from 1 to 1819 A.D., only 5 data points exist. Given this, one would assume better to split the dataset into zones and perform different calibrations. This was tried in the model but it did not yield better results than equation 3 for the whole calibration. In figure 3, the difference between observed data and the one of the model is shown on the right. A close-up of figure 2 is shown for the period 1800-2100 A.D.

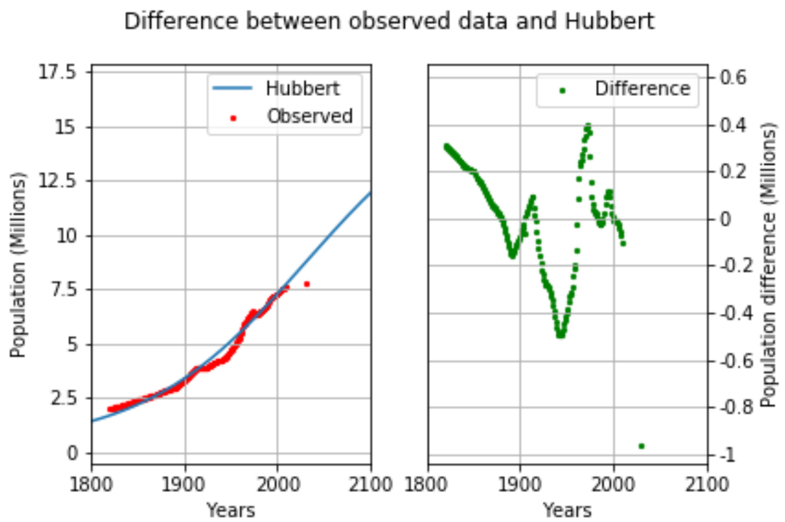

Figure 3: On the left, a close-up of figure 2 showing the projection beyond 2009. On the right, the difference between data points in observed and Hubbert.

References

1. Bacaër, N. (2011). *A short history of mathematical population dynamics*. Springer Science & Business Media.
2. Scherer, L. (2017a). Fundamentals of modelling and data analysis [Powerpoint slides – lecture 3.3]. Retrieved from Brightspace
3. Scherer, L. (2017b). Fundamentals of modelling and data analysis [Powerpoint slides – lecture 5.2]. Retrieved from Brightspace
4. Bolt, J. and J. L. van Zanden (2014). The Maddison Project: collaborative research on historical national accounts. The Economic History Review, 67 (3): 627–651.