

The growth(height) curve analysis by the SITAR model

Interim report for ST606

Yang Wang (19250003)

Contents

1	Introduction	1
2	Methods	1
2.1	The SITAR model	1
3	The results so far	2
3.1	The analysis for the datasets	2
3.2	The first dataset (Boys)	2
3.2.1	Fit the models by using sitar model to find the best model	2
3.2.2	Compare the random effect ($\alpha(a) + \beta(b) + \gamma(c)$)	2
3.2.3	The distance and velocity curves	3
3.3	The second dataset (Girls)	4
3.3.1	Fit the models by using sitar model to find the best model	4
3.3.2	Compare the random effect ($\alpha(a) + \beta(b) + \gamma(c)$)	5
3.3.3	The distance and velocity curves	5
3.4	Compare the growth rate and velocity of height by gender	6
3.5	Conclusion	7
4	Further work	7

1 Introduction

A growth curve is an empirical model of the evolution of a quantity over time. The growth curves are widely used in biology for quantities such as population size or biomass (in population ecology and demography, for population growth analysis), individual body height or biomass (in physiology, for growth analysis of individuals).

Growth is influenced by many factors such as heredity, genetic or congenital factors, illness and medications, nutrition, hormones, and psychosocial environment. A child growth curve is an important standard to help parents to find out whether or not their children follow a normal growth pattern. In fact, hundreds of medical issues can affect children's growth. An abnormal growth pattern is often the earliest indication of an underlying medical condition. Following a normal growth pattern is a good indicator of a child's overall good health.

The growth curve analysis is a statistical issue in life course. Height in puberty involves a growth spurt, the timing and intensity of which varies between individuals.

2 Methods

I used the SuperImposition by Translation And Rotation (SITAR) model, a shape invariant model with a single fitted curve. Curves for individuals are matched to the mean curve by shifting their curve up-down (representing differences in mean size) and left-right (for differences in growth tempo), and the age scale is also shrunk or stretched to indicate how fast time passes in the individual (i.e. velocity). These three parameters per individual are estimated as random effects while fitting the curve. The outcome is a mean curve plus triplets of parameters per individual (size, tempo and velocity) that summarize the individual growth patterns.

2.1 The SITAR model

The method used to summarize the individual growth curves. This is a shape invariant model that involves fitting the following random effects model to the set of height growth curves.

$$y_{it} = \alpha_i + h\left(\frac{t - \beta_i}{\exp(-\gamma_i)}\right)$$

where y_{it} is height for subject i at age t , $h(t)$ is a natural cubic spline curve of height vs age, and α_i , β_i and γ_i are subject-specific random effects. The aim is to choose the values of α_i , β_i and γ_i to make the individual growth curves as similar as possible.

1. α_i is a random height intercept that adjusts for differences in mean height—here it is termed size. Geometrically it can be thought of as a subject-specific shift up or down or translation in the spline curve, with α_i smaller for shorter children.
2. β_i is a random age intercept to adjust for differences in the timing of the pubertal growth spurt in individuals, and it is here called tempo. Geometrically it corresponds to a subject-specific left-right shift or translation in the spline curve, with β_i negative for early puberty and positive for late.
3. γ_i is a random age scaling that adjusts for the duration of the growth spurt in individuals. Its parameterization as $\exp(-\gamma_i)$ ensures that both positive and negative values are permissible, with zero corresponding to average velocity, and the minus sign means that it measures velocity rather than its inverse.

Geometrically it corresponds to a shrinking or stretching of the age scale. Thus for a child with a relatively short spurt and a steep growth curve, γ_i is positive to stretch the age scale and reduce the slope (i.e. peak velocity). Conversely, if γ_i is negative the age scale is shrunk and the curve slope increased. For this reason the parameter is termed velocity.

3 The results so far

3.1 The analysis for the datasets

The file [fit_database_anthropometric_all.xlsx] contains 102,642 records (Individuals: 14,267 (7,239 boys, and 7,028 girls)).

Each data record contains the individual ID, measurement date, age, gender, height, weight, BMI, WHO z-score and WHO z-score categories columns.

Deleted the 7,745 records that the height is NA.

Chose the datasets are heights for boy (31) and girl (38) at the stage between 2007 and 2018 (6-18 year-old), and have the complete set of 22 observations between the ages of 6 and 18 years.

3.2 The first dataset (Boys)

The 31 boys who have the complete set of 22 observations.

3.2.1 Fit the models by using sitar model to find the best model

Although, the $\log(\text{height}) \sim \text{age}$ model is the best, however, it has convergence error. So I chose the $\log(\text{height}) \sim \log(\text{age})$ model.

model 1: $\text{height} \sim \text{age}$

model 2: $\log(\text{height}) \sim \text{age}$

model 3: $\text{height} \sim \log(\text{age})$

model 4: $\log(\text{height}) \sim \log(\text{age})$

```
## Warning in nlme.formula(y ~ fitnlme(x, s1, s2, s3, s4, a, b, c), fixed = s1 + :  
## Iteration 3, LME step: nlminb() did not converge (code = 1). PORT message: false  
## convergence (8)
```

##	formula	variance	residual	AIC	Converge
## 1	height~age	96.56	1.2995116	2718.42	True
## 2	log(height)~age	96.47	0.00854324	-4139.344	false
## 3	height~log(age)	96.45	1.31863340	2735.403	True
## 4	log(height)~log(age)	96.41	0.008600255	-4131.781	True

3.2.2 Compare the random effect ($\alpha(a) + \beta(b) + \gamma(c)$)

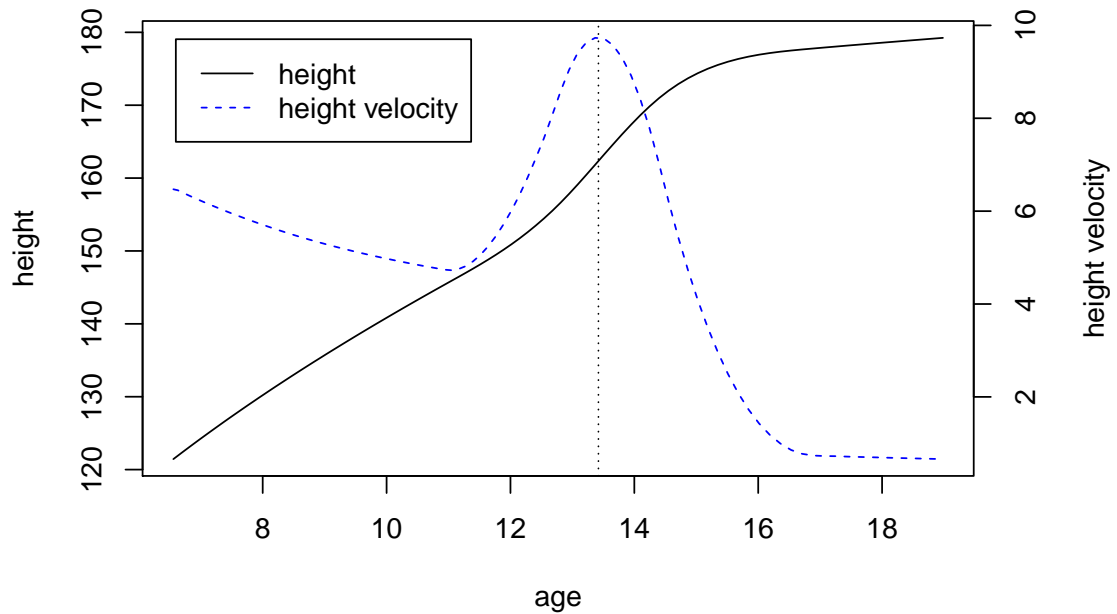
The model with all random effects is best.

##	Type	variance	residual	AIC	BIC
## 1	Random Effect by a	85.27	0.01741566	-3424.986	-3393.311
## 2	Random Effect by a + b	93.48	0.01158098	-3845.512	-3800.262
## 3	Random Effect by a + b + c	96.41	0.008600255	-4131.781	-4068.43

3.2.3 The distance and velocity curves

Distance curve(Black solid line): the amount of height achieved at a given age.

Velocity curve(Blue dash line): the rate of growth at a given age.



```
##      apv      pv
## 13.420  9.739
```

1. The growth rate:

The boys grow fast between 6~16 year-old, grow slowly after 16 year-old.

2. The growth velocity:

- It reached the peak of 9.685 at the age of 13.42.
- It declined slowly between 6~11 year-old.
- It increased rapidly between 11~13.42 year-old.
- It declined rapidly between 13.42~17 year-old.
- It is relatively constant after 17 year-old.

3.3 The second dataset (Girls)

The 38 girls who have the complete set of 22 observations.

3.3.1 Fit the models by using sitar model to find the best model

Although, the $\log(\text{height}) \sim \text{age}$ model is the best, however, it has convergence error. So I chose the $\log(\text{height}) \sim \log(\text{age})$ model.

model 1: $\text{height} \sim \text{age}$

model 2: $\log(\text{height}) \sim \text{age}$

model 3: $\text{height} \sim \log(\text{age})$

model 4: $\log(\text{height}) \sim \log(\text{age})$

```
## Warning in nlme.formula(y ~ fitnlme(x, s1, s2, s3, s4, a, b, c), fixed = s1 + :  
## Iteration 1, LME step: nlminb() did not converge (code = 1). PORT message: false  
## convergence (8)
```

##	formula	variance	residual	AIC	Converge
## 1	height~age	96.68	1.0929107	2930.909	True
## 2	log(height)~age	96.63	0.007419857	-5201.022	False
## 3	height~log(age)	96.30	1.10268133	2942.887	True
## 4	log(height)~log(age)	96.29	0.007452629	-5194.83	True

3.3.2 Compare the random effect ($\alpha(a) + \beta(b) + \gamma(c)$)

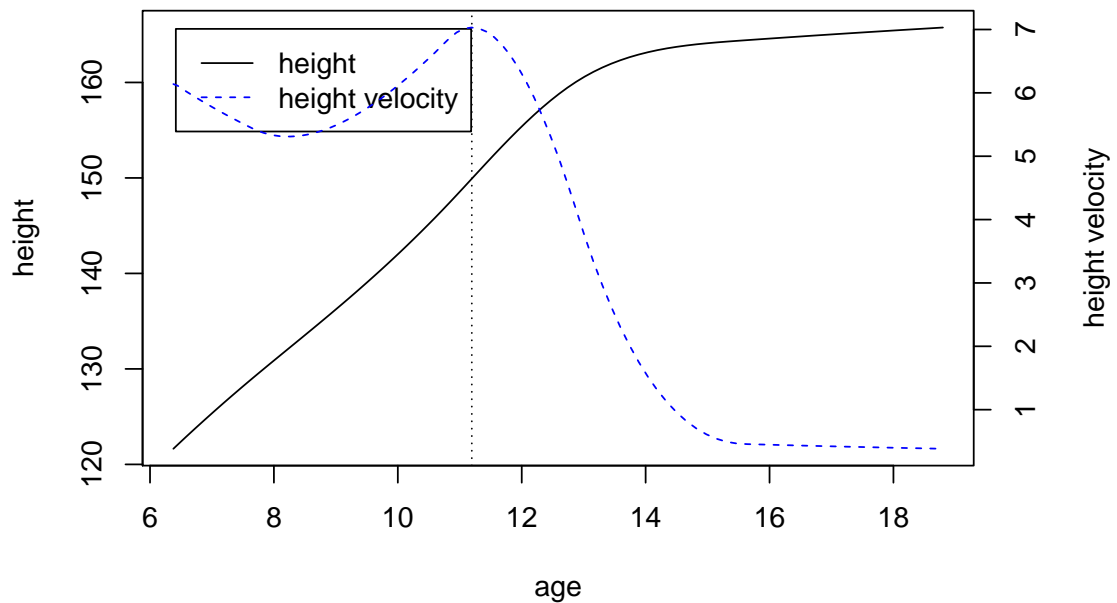
The model with all random effects is best.

##	Type	variance	residual	AIC	BIC
## 1	Random Effect by a	86.45	0.01423831	-4416.925	-4388.713
## 2	Random Effect by a + b	94.27	1.3751748	3195.192	3242.212
## 3	Random Effect by a + b + c	96.29	0.007452629	-5194.83	-5133.704

3.3.3 The distance and velocity curves

Distance curve(Black solid line): the amount of height achieved at a given age.

Velocity curve(Blue dash line): the rate of growth at a given age.



```
##      apv      pv
## 11.190  7.031
```

1. The growth rate:

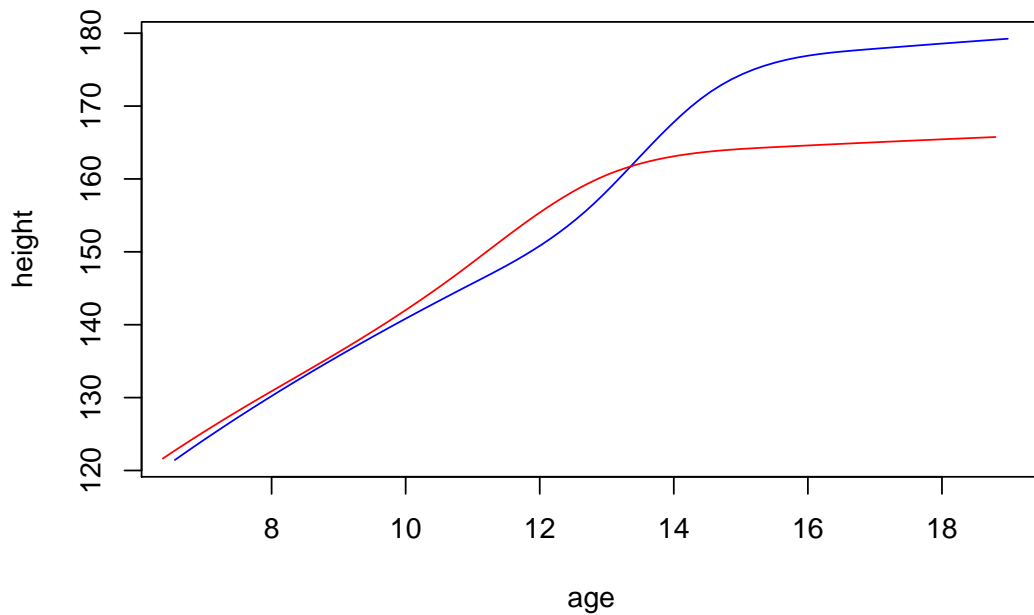
The girls grow fast between 6~14 year-old, grow slowly after 14 year-old.

2. The growth velocity:

- It reached the peak of 7.031 at the age of 11.19.
- It declined between 6~8 year-old.
- It increased between 8~11.19 year-old.
- It declined rapidly between 11.19~15 year-old.
- It is relatively constant after 15 year-old.

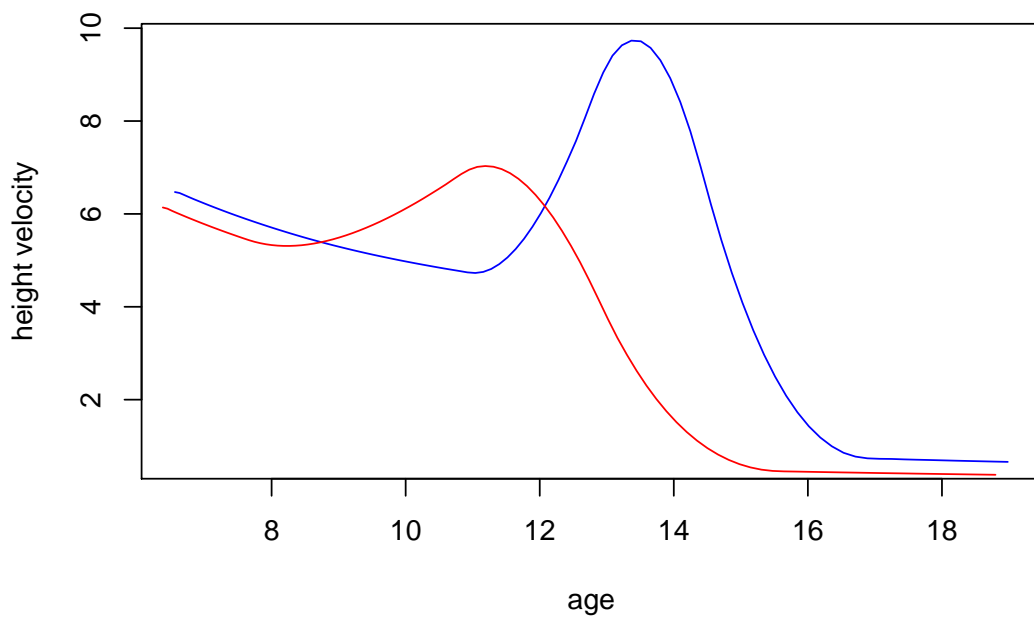
3.4 Compare the growth rate and velocity of height by gender

The Growth Curve (Boy:blue, Girl:red)



1. 6-10 year-old: Boys and Girls almost have the same rate.
2. 10~13.5 year-old: Girls are a little higher than boys.
3. 13.6~18 year-old: Boys are significant higher than Girls.

The Growth Curve : velocity (Boy:blue, Girl:red)



1. 6~9 year-old: Boys and Girls are near to each other and decline.
2. 9~11 year-old: Boys and girls have different trend, Boys continue declining, Girls increase

and reach the top at 11 year-old.

3. 11~15.5 year-old: Girls keep declining fast, Boys increase fast and reach the top at 13.5, then declined fast til 16.5 year-old.
4. The girls keep constant from 15.5 year-old, the boys keep constant from 17 year-old and near to each other.

3.5 Conclusion

The fact that the model fits well both in boys and girls. Although, the log height scale versus age is the best model, however, the convergence error happened when fit the model, so chose the log height scale versus log age scale as the model. Compared the growth curve, we can see that the difference of growth rate and velocity by gender at a certain age.

4 Further work

1. Analyse the results that have done.
2. Try more transformations of model to find a better model.
3. Optimize the parameters of the SITAR model to make the curve better.
4. Focus on Analysing the fixed and random effects how to effect the model.
5. Compare the curves with the standard growth curve (WHO).