Constructing growth reference curves for a cohort of South African children

Melinda Ross

Supervised by Francesca Little

Submitted in partial fulfilment of the requirements of the degree of MSc Biostatistics 2021

Department of Statistical Sciences
University of Cape Town

Contents

List of tables

2 Data

3 Centile derivation methods

4 Derivation of growth curves

4.1 Weight-for-age: boys

4.2 Weight-for-age: girls

Appendix A

List of figures

2 Data

4 Derivation of growth curves

4.1 Weight-for-age: boys

List of abbreviations

AIC	Akaike Information Criterion
BCCG	Box-Cox Cole and Green
BCPE	Box-Cox Power Exponential
BCT	Box-Cox T
BIC	Bayesian Information Criterion
CDF	Cumulative Distribution Function
DCHS	Drakenstein Child Health Study
edf	effective degrees of freedom
FP	Fractional Polynomial
GAIC	Generalised Akaike Information Criterion
GAM	Generalised Additive Model
GAMLSS	Generalised Additive Models for Location, Scale and Shape
GAMM	Generalised Additive Mixed Model
GD	Global Deviance
GLM	Generalised Linear Model
GLMM	Generalised Linear Mixed Model
HIV	Human Immunodeficiency Virus
MGRS	Multicentre Growth Reference Study
NCHS	National Centre for Health Statistics
PDF	Probability Density Function
Q-Q plot	Quantile-Quantile plot
SBC	Shwarz Bayesian Criterion
SD	standard deviation
WHO	World Health Organisation

Acknowledgements

I would like to thank Professor Francesca Little for conceptualising and supervising this dissertation. Her patient guidance and support have been much appreciated and I have been most fortunate to have access to her insight and vast experience over the course of this research.

I would also like to thank my family and friends for believing in me and supporting me along this journey, always keeping up a steady stream of encouragement.

The data used in this dissertation was made available by the Drakenstein Child Health Study. I commend all those involved in ensuring the success of the study, particularly the participants and their families.

And finally, I am very grateful to the National Research Foundation for their financial support.

Abstract

Childhood growth impacts the future welfare of an individual and ultimately the nation. The importance of childhood growth monitoring with growth curves that accurately represent the growth of the population of interest cannot be overemphasised.

This dissertation sought to model the growth of a cohort of South African children and compare their growth to the World Health Organisation (WHO) 2006 Child Growth Standards. Growth reference curves were derived using parametric and semi-parametric methods within the Generalised Additive Models for Location, Scale and Shape (GAMLSS) framework. Various distributions for the growth measurements were compared as well as various curve smoothing approaches for the longitudinal profiles, including cubic splines, fractional polynomials and Berkey-Reed First and Second Order models. The preferred approach was to use the Box-Cox Power Exponential (BCPE) distribution with curve smoothing by cubic splines. Non-parametric quantile regression served as a confirmation that the chosen parametric distributions were appropriate for the data.

A comparison of the derived growth references to the WHO (2006) standards revealed deviations in the patterns of growth and a greater likelihood of diagnosing a child as underweight, stunted or having micro- or macrocephaly when measured against the WHO standards. The poor socioeconomic status and associated harmful exposures of the cohort were noted as potential contributing factors. A fair comparison would require a reasonably healthy and representative sample of the South African population.

These findings do however call into question the appropriateness of the WHO standards for measuring the growth of South African children and bring into focus the value of developing national growth standards.

1 Introduction

1.1 Background

How a child grows and develops physically provides a measure of the health of a population. During the early years of a child's life, the rates of growth and development are at their highest and children are most vulnerable to disease as well as nutritional and stimulation deficiencies. This may have a permanent impact on physical and mental capacity, ultimately impacting the future growth and welfare of a nation (Tarwa and De Villiers, 2007). Childhood growth monitoring is therefore fundamentally important in order to identify signs of growth abnormalities such that appropriate interventions can be implemented timeously.

Growth is monitored using anthropometric indicators based on measures such as weight, height and head circumference. These indicators may be represented graphically with growth curves which are used to monitor the growth of a child by comparing his/her growth trajectory to the population centile curves. Abnormal growth may be suggested by concerning patterns in a child's growth trajectory. These include data points which cross centile lines quickly or values which lie further than two standard deviations from the mean (WHO, 2008).

Growth curves can represent either growth standards or references. Cole (2012) explains that growth standards describe how children should grow: the growth that would be achieved in healthy children living in an environment conducive to them attaining their full genetic growth potential. Deviations from such standards are considered to be evidence of abnormal growth. Cole (2012) further explains that a growth reference, on the other hand, describes how a group of children from a particular region grew over a particular time period and does not necessarily provide a measure of normative growth.

WHO (2006) introduced a new set of growth standards intended for international use and since then these standards have been adopted by more than 100 countries worldwide. The WHO's stance is that children up until age five should all grow similarly if they were breastfed as infants, come from economically advantaged families and growth is unconstrained.

De Onis *et al*. (2004) explain that the impetus for the WHO (2006) Child Growth Standards was an investigation by the WHO into the National Centre for Health Statistics/World Health Organisation (NCHS/WHO) growth references which had been commonly used internationally since the 1970's. Several studies had found that healthy breastfed infants deviate significantly from the NCHS/WHO references. The WHO found a number of shortcomings:

- The population of children used, being from a single area in the USA and of European descent, was not representative of an international population.
- With measurements taken every three months, there was insufficient data to capture the rapid growth in early infancy.
- These infants were also predominantly formula-fed.
- Additionally, the statistical methods available when the references were constructed were not sufficiently complex to capture early childhood growth patterns

(Wang *et al*., 2006; de Onis *et al*., 2004).

To ensure that the study population for the WHO (2006) standards would be representative of the international population, data were collected from six countries, representing the major geographic regions of the world: Brazil, Ghana, India, Norway, Oman and the United States (de Onis *et al*., 2004).

The study consisted of two components: a longitudinal study from birth to 24 months and a cross-sectional study from 18-71 months with an overall sample size of approximately 8440. The purpose of the longitudinal component was to facilitate lactation support in early infancy, assess selection biases, as well as to provide longitudinal measurements to be used to construct growth velocity curves which indicate the rate of growth at a given age. Mother and infant pairs were screened and enrolled at birth and 21 visits were done at weeks 1, 2, 4, 6 and then monthly from two to twelve months and then every two months thereafter for the remainder of the longitudinal study (de Onis *et al*., 2004).

Strict selection criteria were imposed on mother and infant pairs to ensure that there were no significant constraints to growth of the children.

- Infants needed to be born at term,
- Infants should have no health, environmental or economic constraints to growth,
- Mothers should follow feeding guidelines,
- Infants should not be exposed to any significant morbidity.
- Additionally, smoking was included as a strict exclusion criterion owing to its negative impact on the health of both the mother and the child

(de Onis *et al*., 2004).

The WHO growth curves were constructed using the Box-Cox Power Exponential distribution with curve smoothing by cubic splines using the GAMLSS software (de Onis *et al*., 2004). Since the influence of kurtosis in the models was not found to be materially significant, these models simplified to the LMS method by Cole and Green (1992).

Natale and Rajagopalan (2014) emphasise the importance of using growth curves that accurately represent the expected pattern of growth for a population of interest. Failure to do so could result in incorrect diagnoses and consequently poor implementation of healthcare interventions and allocation of resources.

A survey conducted by de Onis *et al*. (2012) found that 125 countries around the world had adopted the WHO (2006) Child Growth Standards by April 2011 and for those that had chosen not to adopt the standards, the main reason cited was that they preferred to use their own national references.

China is one country which endorses their own growth references, last updated in 2009. Yang *et al*. (2015) compared the differences between the WHO (2006) growth standards and the China growth references in assessing childhood malnutrition from birth to five years and found that the prevalence of stunting, underweight and wasting were all higher based on the Chinese references.

There have been numerous studies which have found differences in growth patterns to the WHO (2006) standards despite the children being from economically advantaged environments. Natale and Rajagopalan (2014) carried out a comparison between data from the WHO Multicentre Growth Reference Study (MGRS) and data from studies performed in 55 countries or ethnic groups. Their findings were as follows:

1. For height, most means were within 0.5 standard deviations of the MGRS means, except for Pacific Islanders, the Netherlands and Finland who were above the threshold and India and Saudi Arabia who were below.
2. Weight was found to vary more than height with average weights generally being higher than the MGRS means. Natale and Rajagopalan (2014) attributed this to the WHO promoting slenderness in the interest of health in the midst of the global obesity pandemic. Another plausible explanation is the strict breastfeeding requirements for mothers and infants in the WHO MGRS study as it is a known phenomenon that breastfed children tend to be leaner than their formula-fed counterparts in the early stages of life. In addition, weights at birth were found to vary significantly even among highly developed nations meaning that the WHO standards may not be appropriate for assessing birth weights globally. For example, birth weights in Iceland are much higher than the MGRS means, whereas those in Japan are much lower.
3. The biggest differences were found amongst head circumferences. The mean values for numerous countries and ethnic groups were greater than 0.5 standard deviations above the MGRS means. Among these were Europeans, Asian Indians, Australian Aborigines, Canadian Cree, Japanese and Pacific Islanders.

1.2 Aim

The aim of this dissertation was to model the growth of a cohort of South African children and compare their growth to the WHO (2006) Child Growth Standards. We wished to both describe growth as well as identify abnormal growth.

1.3 Objectives

The objectives of this dissertation were to:

1. assess the growth of a South African cohort in terms of the WHO growth standards,
2. summarise and critically compare different approaches to deriving centiles for growth measurements, including adjustments for longitudinal data,
3. derive growth reference curves for a South African birth cohort and compare them with the WHO standards,
4. identify and compare abnormal growth, specifically stunting, underweight, and micro- and macrocephaly.

2 Data

2.1 Data description

The data used in this dissertation consist of longitudinal anthropometry measures including weight, height and head circumference for children enrolled in the Drakenstein Child Health Study (DCHS) (Zar *et al*., 2015), a multi-year birth cohort study which seeks to understand the impact of environmental, nutritional, psychosocial, maternal, disease and immunological risk factors on child health.

Enrolment occurred prenatally and the mother-infant pairs were followed from birth to beyond five years of age, although the focus of this dissertation is on the first five years. Scheduled visits occurred at birth, between 6 and 10 weeks, 10 weeks, 14 weeks, 6 months, 9 months, 12 months, and then every six months thereafter until 60 months. Additional visits occurred in the event of illness.

The multiethnic study population is from a semi-rural area within the Drakenstein Municipality near Cape Town, South Africa. Zar *et al*. (2015) state that, being of low socioeconomic status, the Drakenstein cohort is subjected to poor living conditions with high levels of exposure to tobacco smoke and alcohol abuse. There is also a high prevalence of malnutrition and infectious diseases such as pneumonia, HIV and tuberculosis. As discussed above, poor socioeconomic conditions, harmful environmental exposures and significant child morbidity are considered constraints to growth by the WHO (2006).

2.2 Data summary

There were 1139 children in the data set, 583 boys and 556 girls. The number of boys and girls that attended each scheduled visit are listed in Table 2.1. 92% of children attended at least three visits and 77% of children attended at least ten. Not all growth measurements were recorded at each visit – the number of weight, height and head circumference measurements recorded per visit for boys and girls are shown in Tables A.1-A.3 in Appendix A. In addition, the median weights, heights and head circumferences for each visit can be seen in Tables A.4, A.5 and A.6 respectively.

Table 2.1: Number of children attending each visit.

Visit	Boys	Girls	Total
Birth	579	553	1132
6-10 weeks	469	426	895
10 weeks	473	439	912
14 weeks	485	454	939
6 months	480	457	937
9 months	435	407	842
12 months	397	386	783
18 months	428	405	833
24 months	383	364	747
30 months	382	368	750
36 months	393	382	775
42 months	440	416	856
48 months	420	406	826
54 months	446	421	867
60 months	422	392	814
Other*	593	399	992

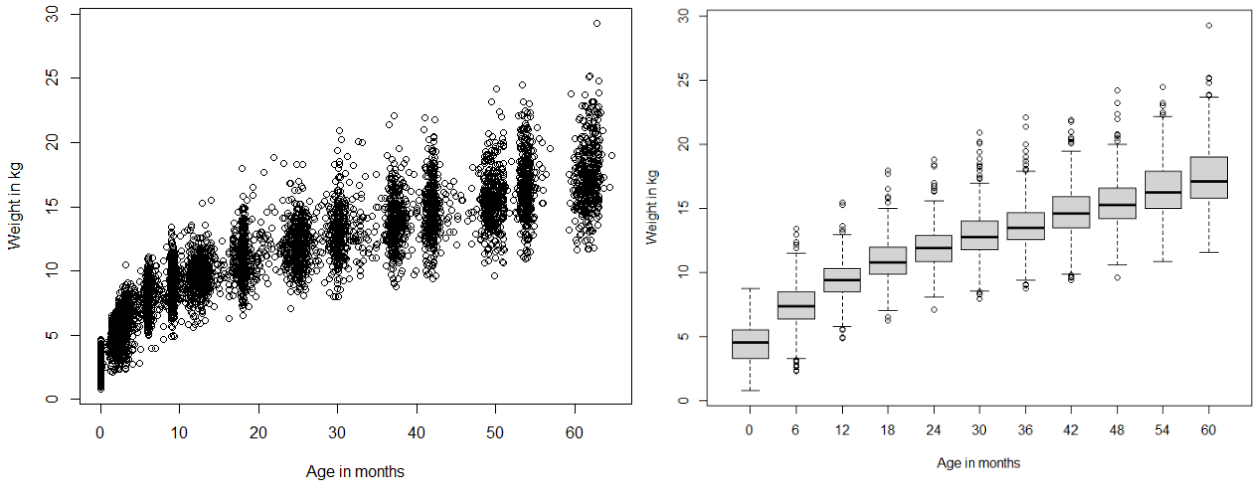
*Other encompasses all ad hoc visits associated with pneumonia episodes.

Plots of the weights, heights and head circumferences vs age as well as boxplots showing the distribution of measurements for each six-month interval from birth are shown for boys and girls seperately in Figures 2.1-2.6. Children initially gain weight at a rapid rate and this rate of increase slows down with increasing age. The fastest weight gain happens in the first year of life. The increasing size of the boxes in the boxplots show that the variance in weights is non-constant and appears to increase with age.

Similarly to weight, the rate of growth in height is initially high in early infancy, slowing to almost linear in the pre-school years. The boxplots suggest that variance may vary slightly with age as well but to a lesser extent than weight.

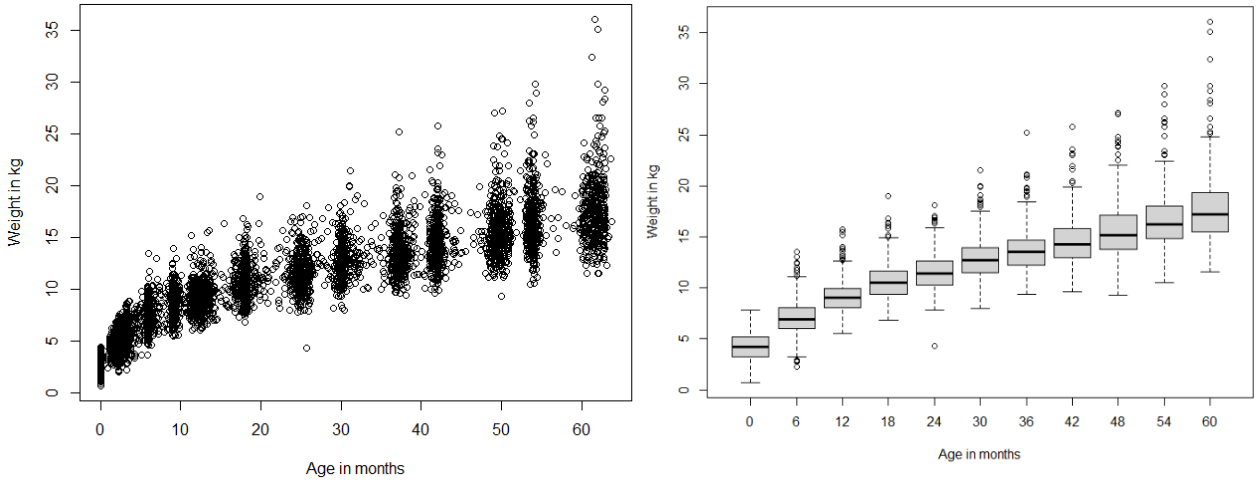
Head circumference grows very rapidly during the first year of life and begins to slow considerably thereafter, particularly after year two. Graber (2021) explains that head

circumference is an indication of brain size and that by 12 months, a child's brain is approximately 75% of adult size. This increases to 80% by age three and 90% by age 7 years. This pattern of growth is evident in Figures 2.5 and 2.6.



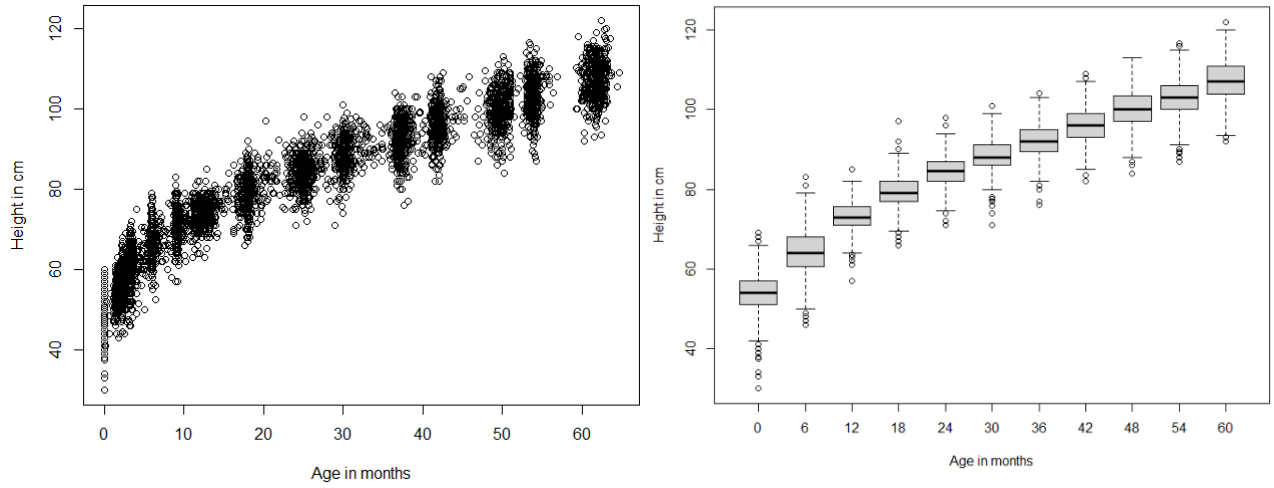
(a) **(b)**

Figure 2.1: (a) Weight vs age for boys and (b) distribution of weights across age for boys.



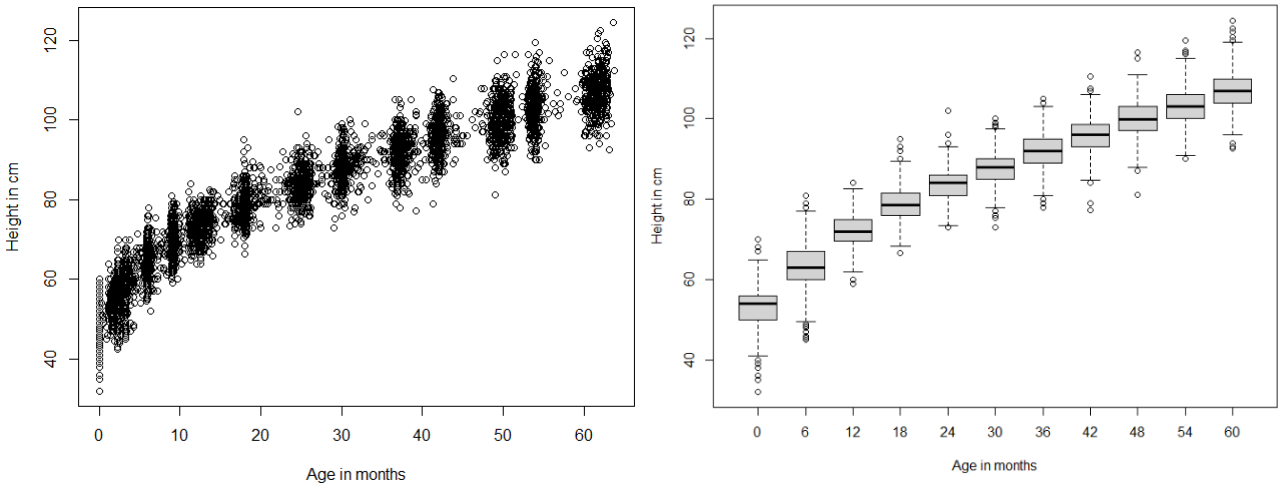
(a) **(b)**

Figure 2.2: (a) Weight vs age for girls and (b) distribution of weights across age for girls.



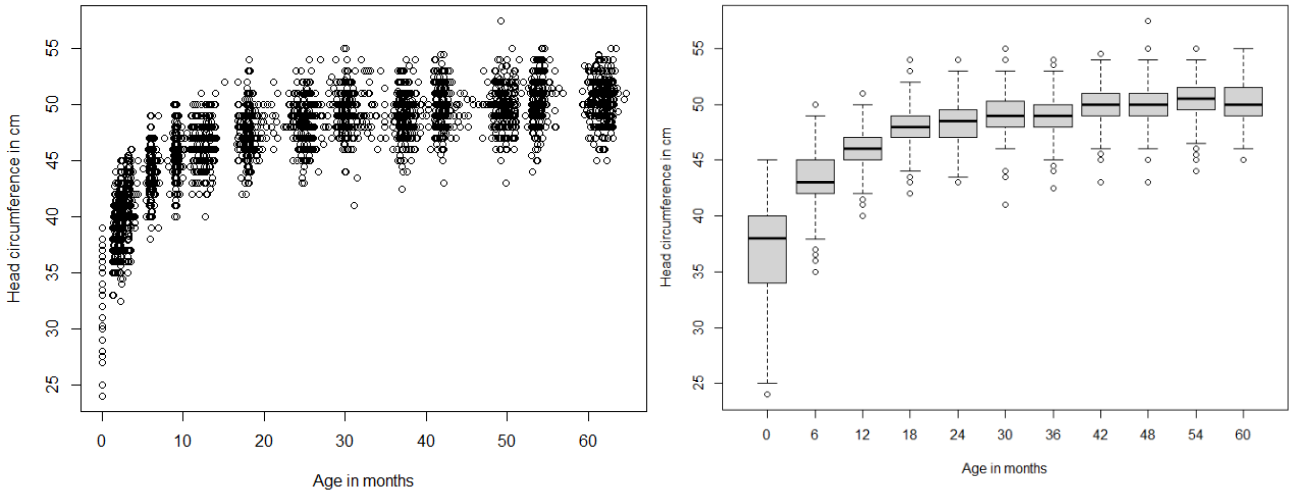
(a) **(b)**

Figure 2.3: (a) Height vs age for boys and (b) distribution of heights across age for boys.



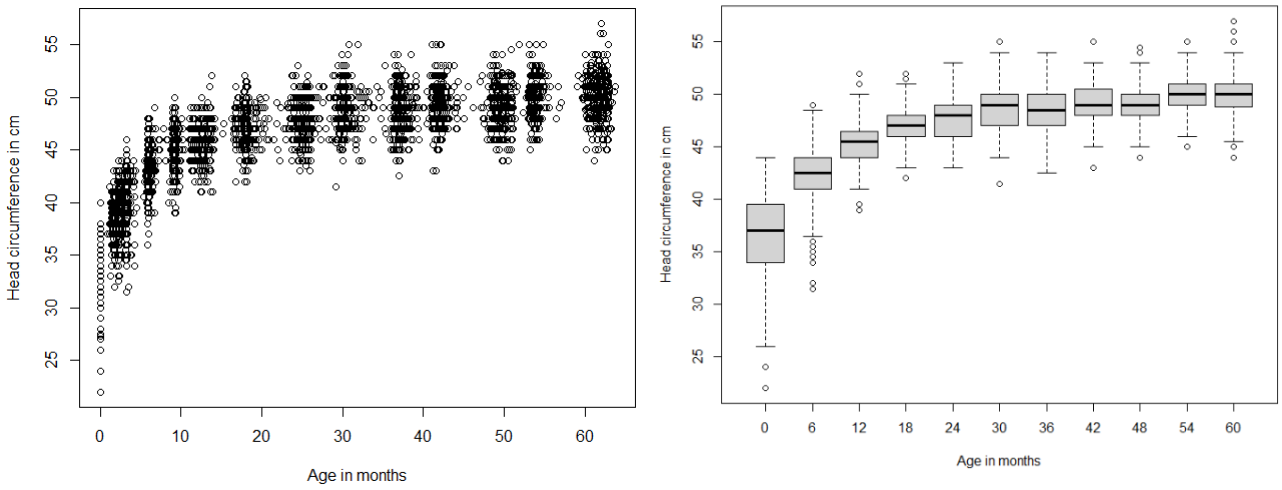
(a) **(b)**

Figure 2.4: (a) Height vs age for girls and (b) distribution of heights across age for girls.



(a) **(b)**

Figure 2.5: (a) Head circumference vs age for boys and (b) distribution of head circumferences across age for boys.



(a) **(b)**

Figure 2.6: (a) Head circumference vs age for girls and (b) distribution of head circumferences across age for girls.

2.3 Assessment in terms of WHO standards

The growth measurements of the children in the DCHS were assessed in terms of the WHO (2006) Child Growth Standards: weight-for-age, length/height-for-age and head circumference-for-age. These standards reflect the growth attained by the child in relation to the expected growth for a given age.

The WHO has made available an R (R Core Team, 2021) package **`igrowup_R`** (WHO, 2006) which contains reference tables for the following standards: weight-for-age, length-for-age, weight-for-length, weight-for-height, BMI-for-age, head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age, and subscapular skinfold-for-age. These tables provided the LMS (described in the methodology section) parameters applicable to each gender and age combination needed to calculate the z-scores, deviations from the expected mean growth in standard deviation (SD) units, for each child at each age they were seen. The reference tables were merged with the DCHS dataset and the z-scores and percentiles were calculated.

The WHO imposed some restrictions on the models which impacted the calculations of centiles and z-scores. Firstly, they restricted the indicators such that the percentiles could be determined only within the interval corresponding to z-scores between -3 and +3 because percentiles beyond +-3 standard deviations are invariant to changes in equivalent z-scores (WHO, 2006).

Furthermore, the weight-related indicators, as well as arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age had distributions that were skewed to the right. With the presence of skewness in the data, the distances between positive z-scores become increasingly large the further away they are from the median. For negative z-scores, the distances become smaller. The Box-Cox Normal distribution used in the LMS method, while capable of modelling skewness, is sensitive to extreme data points. Consequently, the distribution was limited to the interval corresponding to z-scores between -3 and +3 and for z-scores falling outside these boundaries, the standard deviation at each age was set to the distance between ±2 SD and ±3SD. This was in the interest of caution as it avoids making

assertions about the distribution of data outside of the range of observed values (WHO, 2006).

The centiles and z-scores were calculated according to the formulae

$$C_{100\alpha}(t) = M(t)[1 + L(t)S(t)Z_\alpha]^{1/L(t)}, \qquad -3 \le Z_\alpha \le 3$$

and

$$z_{ind} = \frac{\left[y/M(t)\right]^{L(t)} - 1}{S(t)L(t)}$$

$$z_{ind} = \begin{cases} z_{ind}, & |z_{ind}| \le 3 \\ 3 + \left(\frac{y - SD3\ pos}{SD23\ pos}\right), & z_{ind} > 3 \\ -3 + \left(\frac{y - SD3\ neg}{SD23\ neg}\right), & z_{ind} < -3 \end{cases}$$

where $SD3\ pos$ is the 3 standard deviation cut-off calculated at time *t*:

$$SD3\ pos = M(t)[1 + L(t) * S(t) * (3)]^{1/L(t)}$$

$SD3\ neg$ is the 3 standard deviation cut-off calculated at time *t*:

$$SD3\ neg = M(t)[1 + L(t) * S(t) * (-3)]^{1/L(t)}$$

$SD23\ pos$ is the difference between the 2 and 3 standard deviation cut-offs calculated at time *t*:

$$SD23\ pos = M(t)[1 + L(t) * S(t) * (3)]^{1/L(t)} - M(t)[1 + L(t) * S(t) * (2)]^{1/L(t)}$$

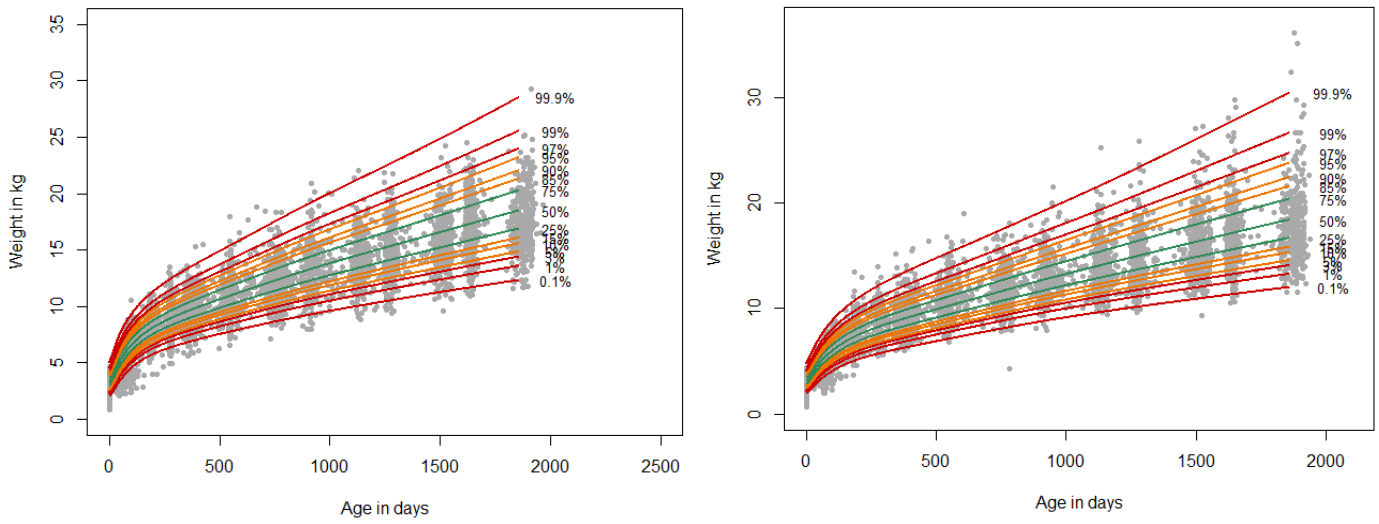
$SD23\ neg$ is the difference between the -2 and -3 standard deviation cut-offs calculated at time *t*:

$$SD23\ neg = M(t)[1 + L(t) * S(t) * (-2)]^{1/L(t)} - M(t)[1 + L(t) * S(t) * (-3)]^{1/L(t)}$$

(WHO, 2006).

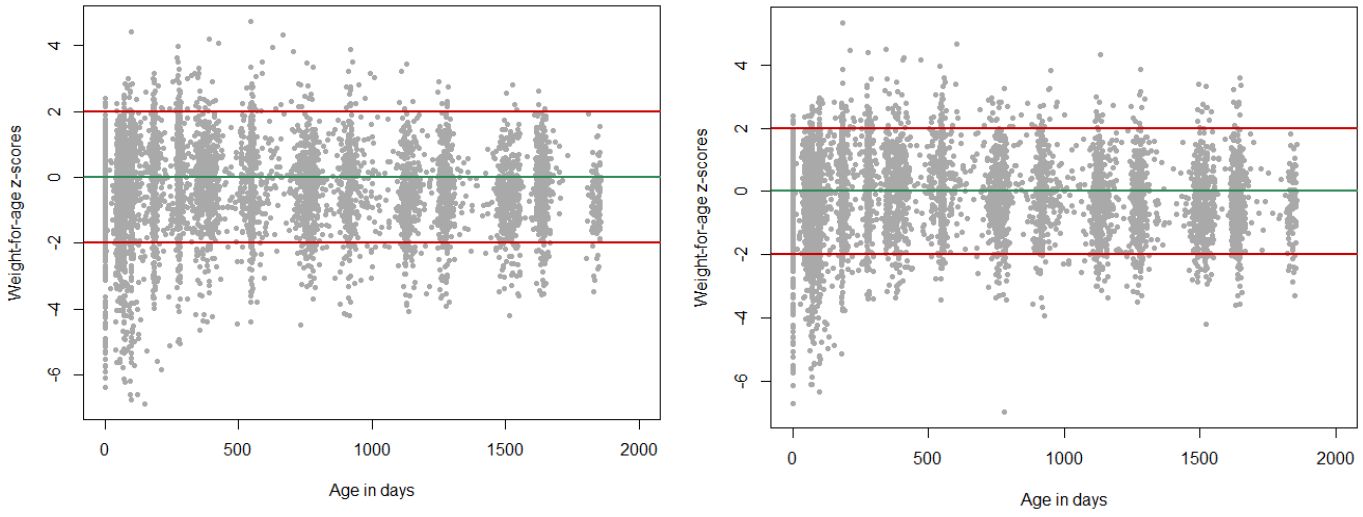
Figure 2.7 shows the observed weights for boys and girls plotted against age with the WHO centile curves overlaid. In addition, the z-scores for weight in terms of the WHO standards and boxplots of these z-scores are shown in Figures 2.8 and 2.9 respectively. The plots for height and head circumference follow in Figures 2.10-2.15.

For the majority of the children's weight measurements, the corresponding z-scores were between -2 and +2. Measurements outside of this range are classified as abnormal growth. There are a large number of children with very low weight-for-age z-scores at birth and for the visits before six months resulting in the average birthweight falling below the 50th centile. The weights then catch up until approximately 18 months, whereafter they once again lag the 50th centile of the WHO standards.



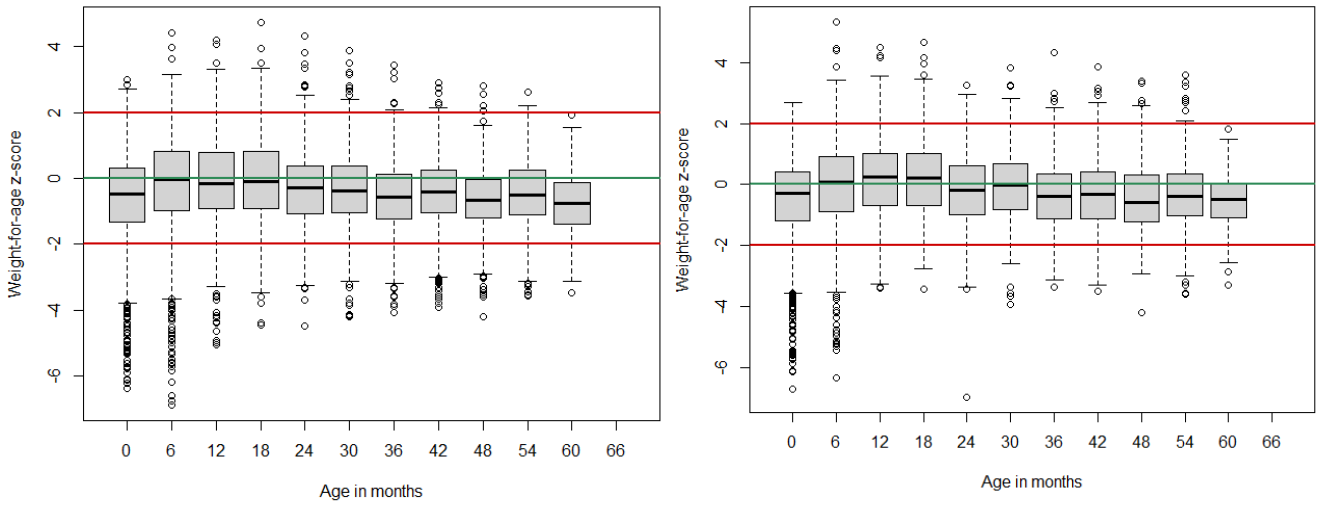
(a) **(b)**

Figure 2.7: Weight vs age with WHO standards overlaid for (a) boys and (b) girls.



(a) **(b)**

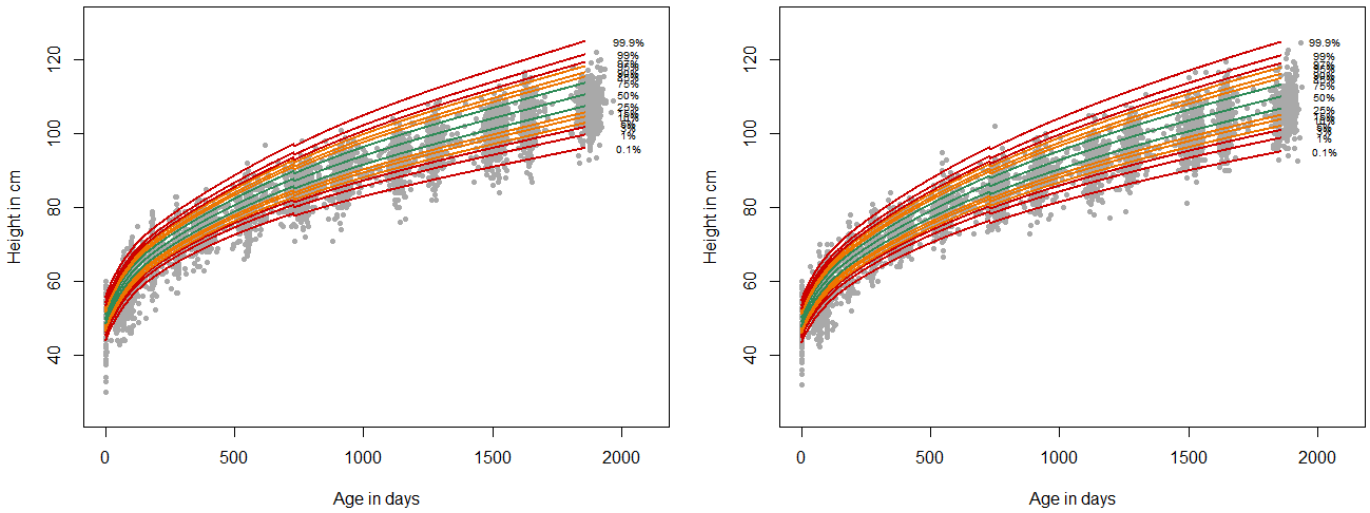
Figure 2.8: Weight-for-age z-scores in terms of WHO standards for (a) boys and (b) girls.



(a) **(b)**

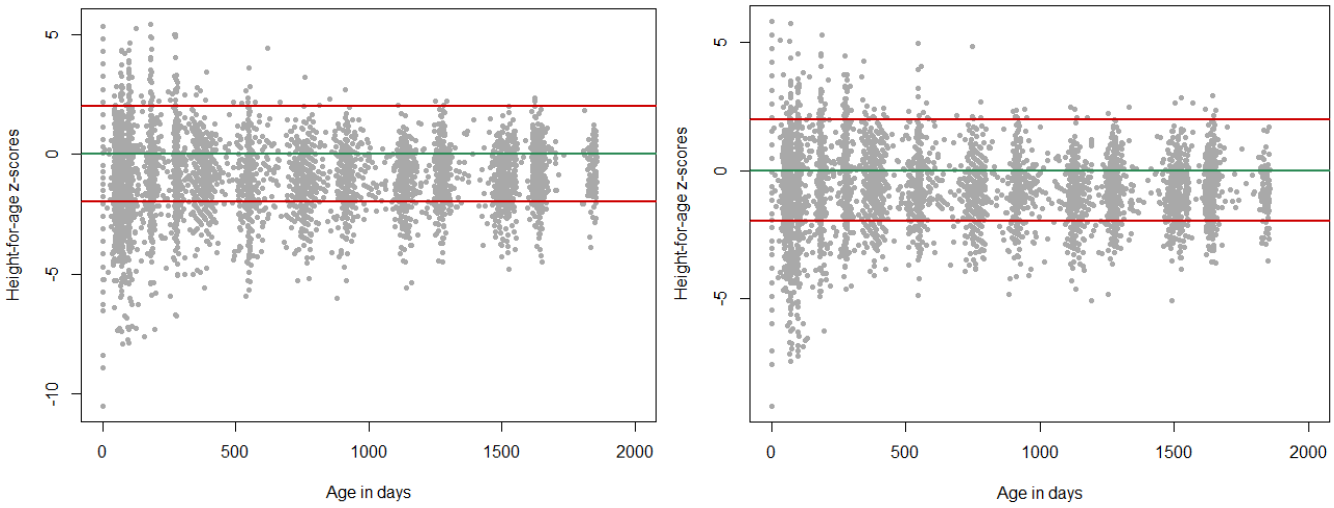
Figure 2.9: Boxplots of weight-for-age z-scores in terms of WHO standards for (a) boys and (b) girls.

The median z-scores for heights all fall well below the WHO 50th centile across all ages. This is particularly pronounced between 18 months and five years of age.



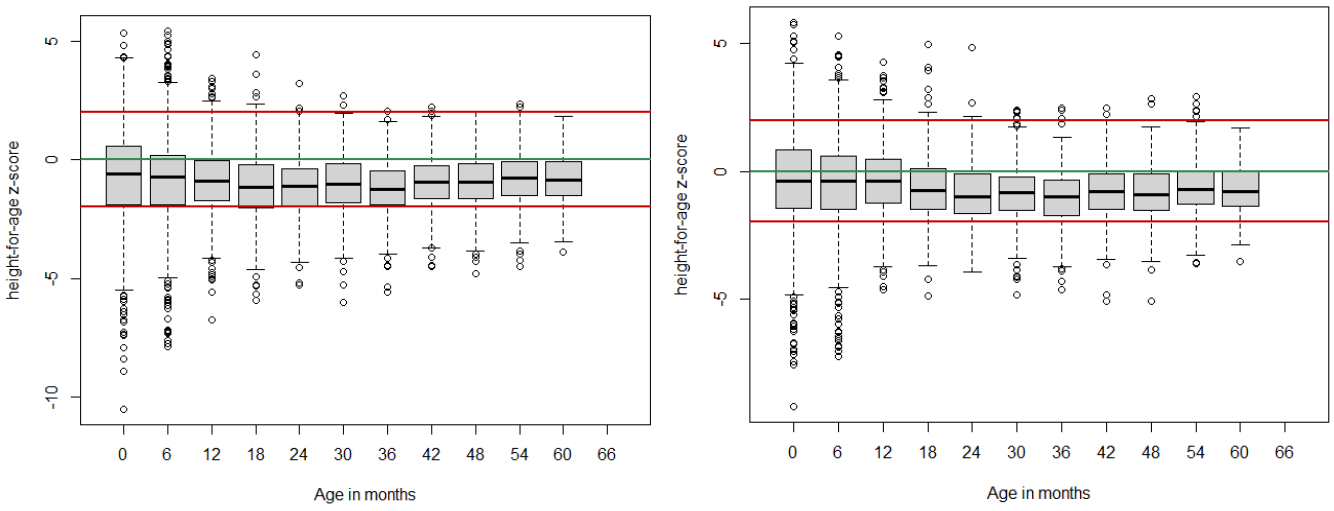
(a) **(b)**

Figure 2.10: Height vs age with WHO standards overlaid for (a) boys and (b) girls.



(a) **(b)**

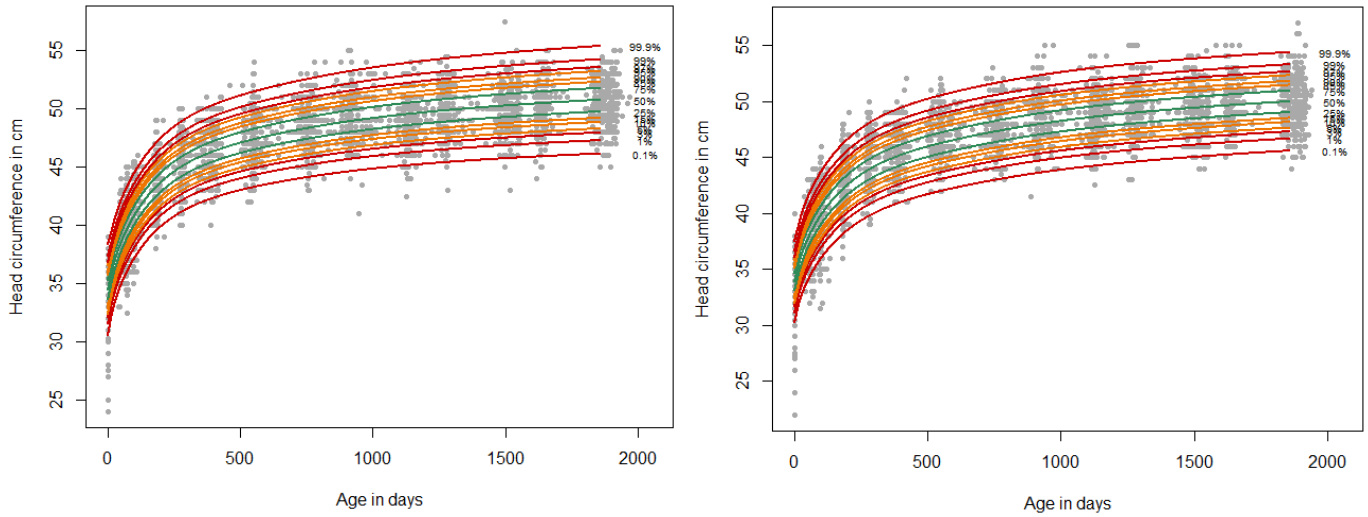
Figure 2.11: Height-for-age z-scores in terms of WHO standards for (a) boys and (b) girls.



(a) **(b)**

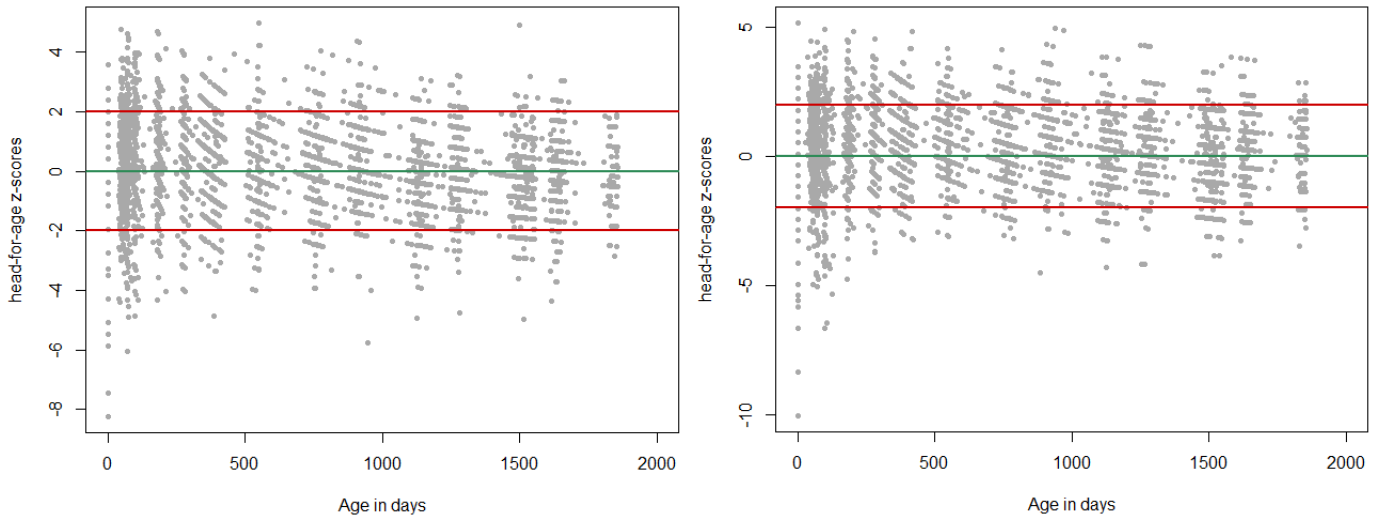
Figure 2.12: Boxplots of height-for-age z-scores in terms of WHO standards for (a) boys and (b) girls.

The median head circumferences for boys and girls exceed the WHO 50th centile during the first two years of life and thereafter are relatively on-par.



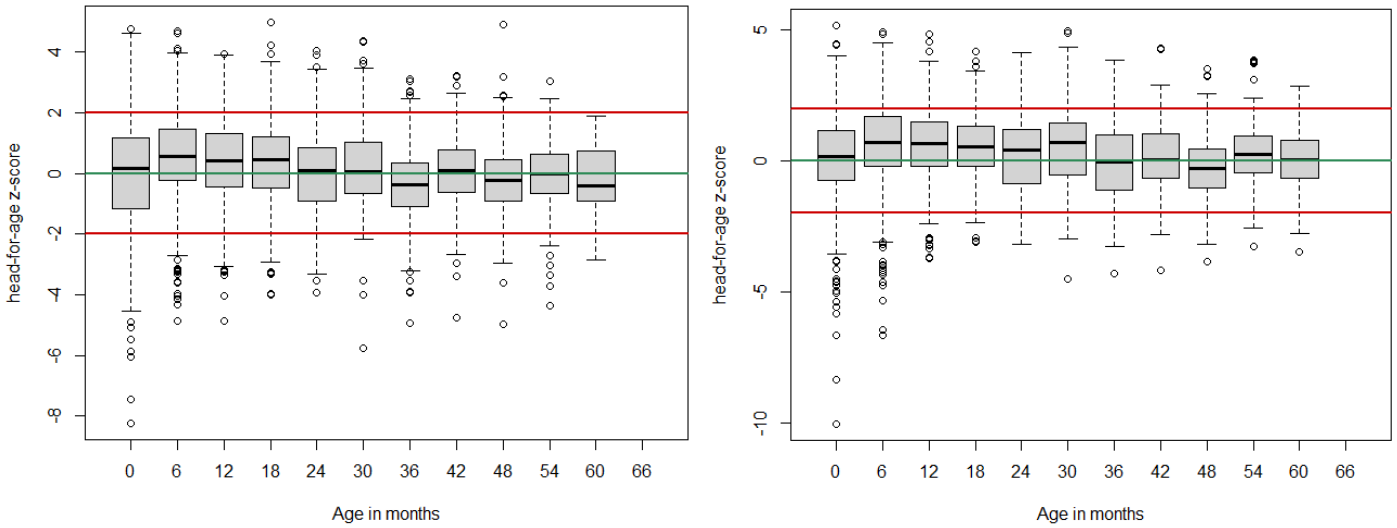
(a) **(b)**

Figure 2.13: Head circumference vs age with WHO standards overlaid for (a) boys and (b) girls.



(a) **(b)**

Figure 2.14: Head circumference-for-age z-scores in terms of WHO standards for (a) boys and (b) girls.



(a) **(b)**

Figure 2.15: Boxplots of head circumference-for-age z-scores in terms of WHO standards for (a) boys and (b) girls.

For all three measures, weight, height and head circumference, differences in the patterns of growth between the Drakenstein cohort and the WHO standards are evident. The biggest differences were observed for height where the median heights were consistently below the WHO 50th centile. Although not as extreme, material differences were also observed for the weight and head circumference measurements. These were not found to be consistently higher or lower than the WHO standards but rather were either lower or higher during certain age ranges which suggests that the children may have different periods of slower and accelerated growth compared to the WHO cohort.

For further comparison, growth reference centiles were derived for the Drakenstein cohort and a discussion and comparison of these methods follows.

3 Centile derivation methods

3.1 Introduction

The lines on a growth chart are called centiles and these show how the growth measurements are expected to vary with age. A particular centile curve represents the line below which a certain percentage of children's measurements lie. For example, the 50th centile for weight means that 50% of children will be at or below that weight at any given age. There are a number of modelling approaches for the construction of growth centiles. These may be grouped into three main categories:

- **Parametric methods**

  These are based on assumptions about the distribution of the data and model the growth pattern with age using a parametric model. Parametric methods are a popular choice as the ease of calculation of z-scores and centiles is desirable. They usually summarise the data using summary statistics based on a combination of the four moments: mean, standard deviation, skewness and kurtosis. If using one of these methods, it is important to carefully assess the distributional properties of the anthropometric measures of interest. The selected model should adequately represent the complexity of the relationship between the growth measurements and age as well as account for non-constant variance and skewness and kurtosis if applicable (Stasinopoulos *et al*., 2017). Borghi *et al*. (2006) recommend studying the empirical median, standard deviation, skewness and kurtosis of the sample across the age range. Borghi *et al*. (2006) also cautioned about the difficulty involved in modelling kurtosis through extrapolation since the information available at the tails is scant. It is recommended that a more complex model involving kurtosis only be chosen if it results in a significant improvement over models based on the first three moments only.

- **Semi-parametric methods**

  Semi-parametric methods involve the assumption of a particular distribution structure, like parametric models, together with the application of non-parametric smoothing methods. Selecting an appropriate degree of smoothing forms an important part of the modelling process. Too little smoothing, and the result will be irregular growth curves

because sampling variability across ages, the effect of outliers and measurement variability will not be smoothed out. Too much smoothing, and the underlying growth pattern may be misrepresented owing to the flattening of peaks and troughs of biological growth. The combination of the distributional assumptions and smoothing technique should aim to represent the underlying biological growth patterns as accurately as possible while also striking a balance with simplicity.

- **Non-parametric methods**
  Non-parametric methods make no distributional assumptions about the data. Quantile regression is an example of such an approach which models the structure of the profile across ages using cubic splines, which may be constructed from a linear combination of basis splines.

3.2 Longitudinal data use in deriving growth centiles

The conventional methods for constructing growth reference curves are based on the use of a cross-sectional data design. The research in this dissertation is based on a longitudinal dataset as the measurements were collected for each child over a series of repeat visits from birth to five years of age. The main purpose of collecting longitudinal data is usually to quantify variation in growth rather than size. This is done using growth velocity curves which are conditional centile curves as they are based on a previously attained size for age. In essence, they serve as individualised growth references.

Royston and Altman (1995) present two approaches to dealing with longitudinal data in growth curve estimation that have been implemented in previous studies and discuss why they consider these two approaches to be incorrect.

The first and simplest of these is to simply ignore the serial nature of the data and estimate the size centiles as though the data were cross-sectional. Altman and Chitty (1994) explain that this is suboptimal and, owing to the non-independence of measurements for the same subject, the effective sample size will be closer to the number of subjects rather than the number of measurements. Despite the shortcomings of this approach, it has been used to estimate reference centiles. Persson (1988) applied it to the estimation of fetal growth curves.

More recently, the WHO (2006) ignored the serial nature of the longitudinal component of the data when deriving the updated growth standards. The WHO MGRS consists of a longitudinal component which extends from birth to 24 months and a cross-sectional component from 18-71 months with an overlap period between 17 and 24 months inclusive. The reason for the longitudinal component was to facilitate the derivation of growth velocity curves. When calculating the size centiles, the longitudinal and cross-sectional datasets were merged and the lack of independence among measurements on the same subject was not accounted for. Borghi *et al*. (2006) explained that the centiles are still considered to be statistically valid.

They argue that, when using methods that apply to cross-sectional data on a longitudinal dataset, as long as the timing of repeat visits is fixed and the number of measurements per child is the same aside from some missing visits, we do not expect the estimate of the mean curve to be biased. Ahuma and Altman (cited in Cole, 2019) reinforce this by stating that estimates from longitudinal data will be unbiased if there is independence between the timing of repeated measures and the actual measures. Borghi *et al*. (2006) caution that the standard errors of the centile estimates are however likely to be affected which will impact the accuracy of the confidence intervals around the centiles. The proposed remedy is to use the bootstrap technique to derive the standard errors associated with the centile estimates. This may be done by resampling children to create new samples of the same size in the bootstrap procedure.

A more sophisticated, but also incorrect, approach for deriving size centiles from longitudinal data is the individual regression lines method. This method was used by Gallivan *et al*. (1993) to estimate centiles for abdominal circumferences and estimated fetal weight. They fitted 67 individual quadratic curves, obtained fitted values for each fetus, calculated the mean and standard deviations of these fitted values and used them to calculate the centiles. This is not a parsimonious approach as a large number of regression coefficients need to be estimated. It also does not take into account variation within each subject with the result that the centiles will be too close together and thus unreliable. Further concerns are that the centiles will not change smoothly with age and each subject needs to have at least three measurements to be included in the analysis. This same method was also used by Deter *et al*. (1982).

Royston (1995) provided a solution to the above problems by showing how multilevel modelling could be used to estimate both conditional growth centiles as well as valid size centiles from longitudinal data. He demonstrated this by modelling the relationship between fetal size and gestational age using fractional polynomials with the incorporation of random effects to account for the correlation amongst measurements from the same fetus. This method was found to produce a parsimonious and reliable solution.

Mixed effect models are well-suited to modelling growth as they allow the variance to be broken down into the contributions of between and within-subject variability. Complex correlation structures may be incorporated as well as dependencies on covariates such as age. The inclusion of non-parametric components is also possible (Borghi *et al*., 2006).
The multi-level modelling approach to calculating size centiles from longitudinal data has since been adopted in numerous studies. Verburg *et al*. (2008) derived reference curves for normal fetal growth based on a number of measures (biparietal diameter, head circumference, transverse cerebellar diameter, abdominal circumference and femur length). They used fractional polynomials to model the relationship between fetal growth and gestational age and also fitted regression lines for the dependency of the residual standard deviation on gestational age. The reference centiles could then be calculated and the curves plotted. Tilling *et al*. (2014) used multilevel fractional polynomial and linear spline modelling to derive centiles for children's height from birth to age 10. They assumed an unstructured variance/covariance matrix and therefore estimated all of the variances and covariances for the random effects.

Wade and Ades (1998) investigated the impact of explicitly modelling within-subject correlations on growth curves fitted using an LMS-based maximum likelihood method. They experimented with various different correlation structures and, although there was a strong correlation structure in the data, it was found that incorporating a correlation structure into the likelihood function did not significantly impact the choice of model for the mean curve or the fitted centiles. In addition, the degree to which the precision of the estimated centiles was impacted when ignoring the serial nature of the data was not deemed to be clinically significant.

3.3 Parametric and Semi-Parametric Methods

3.3.1 Fractional Polynomials

Polynomials may be used when the relationship between the response variable and its predictors is a curved one. This may be addressed through the addition of a quadratic term and sometimes cubic and higher order terms when necessary. Royston and Altman (1994) highlight that low-order polynomials do not offer many curve shapes and therefore often do not fit the data adequately. Higher order polynomials, while they are able to follow the data more closely, often do not fit the data well at the extremes. To overcome these limitations, Royston and Altman (1994) introduced fractional polynomials.

Fractional polynomial models are a popular choice for modelling childhood growth as they offer the flexibility of a wide range of curve shapes while remaining parsimonious. They are able to produce comparable model fits to conventional polynomials but with fewer terms and have the added advantage of the ability to approximate asymptotes. Fractional polynomials were used by Wen *et al*. (2012) to model BMI. Tilling *et al*. (2014) used fractional polynomials and a multilevel modelling approach to model childhood growth in terms of height from birth to age 10. Their best fitting fractional polynomial was one with powers of the square root of age and the square root of age multiplied by the log of age.

A fractional polynomial of degree *m* is given by the formula

$$\phi_m(X;\, \xi, p) = \xi_0 + \sum_{j=1}^{m} \xi_j X^{(p_j)}$$

where *m* is a positive integer,
$p = (p_1, \ldots, p_m)$ is a real-valued vector of powers with $p_1 < \cdots < p_m$,
$\xi = (\xi_0, \ldots, \xi_m)$ are real-valued coefficients
and *X* represents an explanatory variable such as age
(Royston and Altman, 1994).

The round bracket notation signifies the Box-Tidwell transformation, that is:

$$X^{(p_j)} = \begin{cases} X^{p_j} & if\ p_j \neq 0 \\ lnX & if\ p_j = 0 \end{cases}.$$

In the case where two powers are the same, one of them is multiplied by lnX. For example, consider a two-term fractional polynomial where $p_1 = p_2$:

$$\phi_2(X;\ \xi, p) = \ \xi_0 + \ \xi_1 X^{(p_1)} + \ \xi_2 X^{(p_2)}$$
$$= \ \xi_0 + \ (\xi_1 + \xi_2)\ X^{(p_1)}\ .$$

This is a fractional polynomial of degree 1 and the limit as p_2 tends to p_1 of

$$\xi_0 + \ \xi_1 X^{(p_1)} + \ \xi_2 X^{(p_1)}(X^{(p_2 - p_1)} - 1)/(p_2 - p_1)$$

is

$$\xi_0 + \ \xi_1 X^{(p_1)} + \xi_2 X^{(p_1)} ln(X)\ .$$

For $m > 2$ and $p_1 = \cdots = \ p_m$, the generalised formula is given by

$$\xi_0 + \ \xi_1 X^{(p_1)} + \ \sum_{j=2}^{m} \xi_j X^{(p_1)}\ (lnX)^{j-1}.$$

Thus, Royston and Altman (1994) present the full definition of a fractional polynomial as given below.

For arbitrary powers $p_1 \ < \cdots < \ p_m$, set $H_0(X) = 1, p_0 = 0$. Then,

$$\phi_m(X;\ \xi, p) = \sum_{j=0}^{m} \xi_j H_j(X)$$

where $j = 1, 2, \ldots, m$

$$H_j(X) = \begin{cases} X^{(p_j)} & if\ p_j \ \neq \ p_{j-1} \\ H_{j-1}(X) lnX & if\ p_j = \ p_{j-1} \end{cases}.$$

Fractional polynomials require that $X>0$ so if non-positive values are present then X needs to be transformed to ensure positivity. One approach is to select a non-zero origin $\zeta < X$ and rewrite the fractional polynomial definition as:

$$\phi_m(X;\ \xi, p) = \sum_{j=0}^{m} \xi_j H_j(X - \zeta).$$

Royston and Altman (1994) recommend choosing powers with replacement from the set $\{-2, -1, -0.5, 0, 0.5, 1, 2, \max(3, m)\}$ with the power of 0 interpreted as lnX.

Choosing the order of the fractional polynomial may be done informally perhaps based on prior information available or by increasing m until there is no further benefit achieved by doing so. Fractional polynomials of order two are usually sufficient for most applications (Royston and Altman, 1994).

Tilling *et al*. (2014) summarise the process of selecting the best fitting fractional polynomial of order one or two as follows:

(1) Choose the best fitting single fractional polynomial by fitting models for the powers *-2, -1, -0.5, 0, 0.5, 1, 2, 3* and comparing them based on the model deviance.
(2) Try all possible combinations of pairs of these powers and compare them based on their model deviances. This gives the best fitting fractional polynomial of degree *2*.
(3) Calculate the difference in model deviances for the best fitting fractional polynomials with one and two terms and compare it to the Chi square distribution with *2* degrees of freedom (since one extra power and coefficient are being estimated) to assess whether the additional term improves the model significantly.

3.3.2 Berkey-Reed Models

The Berkey-Reed model is a simple linear model which was proposed by Berkey and Reed (1987) for modelling childhood growth in terms of length and potentially also weight and head circumference. Its development was prompted by the fact that other commonly used linear models for growth have been shown to suffer from deficiencies. It is well suited to modelling childhood growth as it incorporates an inflection point which may capture changes in growth velocity. Berkey and Reed (1987) explain that the normal growth pattern (rapid growth in early infancy which becomes more gradual with increasing age until almost linear growth is attained during the pre-school years) may be interrupted by illness or environmental and nutritional issues which could impact the rate of growth.

Two forms of the model were proposed: a First Order model and a Second Order model. The First Order model for a response variable *Y* and explanatory variable *X (age)* is given by:

$$y_{ij} = \beta_{0i} + \beta_{1i}x_{ij} + \beta_{2i}\ln(x_{ij}) + \beta_{3i}\frac{1}{x_{ij}} + \varepsilon_{ij},$$

where y_{ij} is the growth response for subject *i* at age *j*,

ε_{ij} are the random measurement errors at each age

and β_i are the parameter coefficients

(Berkey and Reed, 1987).

An important advantage of the Berkey-Reed model is that it is a structural model with parameters that have biological interpretation.

- β_{0i} is related to the growth measurement at birth
- β_{1i} is related to the linear component of the growth velocity
- β_{2i} is related to the deceleration in growth velocity
- β_{3i} represents the inflection point

The Berkey-Reed Second Order model is an extension of the First Order model which incorporates a second inflection point and is given by:

$$y_{ij} = \beta_{0i} + \beta_{1i}x_{ij} + \beta_{2i}\ln(x_{ij}) + \beta_{3i}\frac{1}{x_{ij}} + \beta_{4i}\frac{1}{{x_{ij}}^2} + \varepsilon_{ij}$$

(Berkey and Reed, 1987).

3.3.3 The GAMLSS Framework

The simple linear regression model for a response variable Yi is of the form $Yi = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \varepsilon_i$. The error terms ε_i are assumed to be independently and identically distributed with a mean of zero and constant variance. An additional assumption imposed by the least squares solution is that the error terms, and therefore also the response variable, are normally distributed. While this model is very straightforward and widely used, the assumptions are often not met in real-world datasets (Stasinopoulos *et al*., 2017).

Linear mixed models extend the simple linear regression model by allowing both fixed and random effects in order to account for non-independence in the data, such as when dealing with repeated measures data or data with a hierarchical structure. The conditional distribution of *y* given the random effects is assumed to follow a normal distribution, which excludes explicit modelling of skewness and kurtosis.

The generalised linear models (GLM) and generalised additive models (GAM) address some of the limitations of the simple linear regression model. These models assume that the response variable *y* follows a distribution from the exponential family and that the mean of Y_i, μ_i, is related to the predictors via a monotonic link function. A GAM, introduced by Hastie and Tibshirani (1986), extends the GLM by modelling the relationships between the dependent variable and the predictor variables using smooth non-parametric functions, allowing for the fitting of semi-parametric models. Since the exponential family comprises a number of distributions, GLMs and GAMs are fairly flexible. However, the variance of the response $V(y) = \phi v(\mu)$ is a function of a dispersion parameter and the mean through the variance function $v(\mu)$ which will not accurately model relationships between the variance of *Y* and the predictors which are different to those described by the variance function. Skewness and kurtosis may also not be accurately accounted for as they are generally functions of the mean and dispersion parameter as well for the exponential family of distributions (Stasinopoulos *et al*., 2017).

Generalised linear mixed models (GLMM) and generalised additive mixed models (GAMM) extend the GLM and GAM through the addition of random effects to account for within subject association due to repeated measurements for the same subject. While these models

are more flexible than GLM and GAM, they still do not allow for the modelling of skewness and kurtosis as functions of the explanatory variables as is possible for the mean. Furthermore, fitting of GLMM's and GAMM's is computationally intensive (Stasinopoulos *et al*., 2017).

The generalised additive models for location, scale and shape (GAMLSS) offer further modelling capability by relaxing the assumption that the response must follow an exponential family distribution and allowing it to follow a more general distribution family. Additionally, all parameters of the distribution may be modelled as functions of the explanatory variables and random effects (Stasinopoulos *et al*., 2017).

For a GAMLSS, it is assumed that the responses $y_i \ , i = 1, 2, \dots, n$ are independent conditional on θ_i, the set of parameters defining the moments of the underlying distribution of the responses. The probability density function is given by:

$$f(y_i|\, \theta_i)$$

where $\theta = (\theta_{i1}, \theta_{i2}, \ \dots \ , \theta_{ip})^T$ is a vector of parameters where each parameter can be related to the explanatory variables and random effects through a monotonic link function $g_k(.)$, given by

$$g_k(\theta_k) = \ \eta_k = \ \boldsymbol{X}_k\beta_k + \sum_{j=1}^{J_k} \boldsymbol{Z}_{jk}\gamma_{jk}$$

where θ_k and η_k are vectors of length *n*,
β_k is a vector of parameters of length J_k',
$\boldsymbol{X}_k$ is a known design matrix of order $n \ \times \ J_k'$,
$\boldsymbol{Z}_{jk}$ is a fixed known $n \ \times \ q_{jk}$ design matrix, and
γ_{jk} is a q_{jk}-dimensional random variable
(Stasinopoulos *et al*., 2017).

The parameters θ_1 and θ_2 usually represent the location (mean) and scale (variance) parameters of the distribution and may be denoted by μ and σ respectively. The remaining parameters represent additional shape parameters where applicable, however the model may be adjusted to the parameters of any population distribution. For most distributions, no more than two shape parameters are needed and may be denoted by $\theta_3 = \ \nu$ and $\theta_4 = \ \tau$, which

usually measure skewness and kurtosis. Thus, a GAMLSS model for a distribution with four parameters $f(y|\mu,\sigma,\nu,\tau)$ has the form

$$g_1(\mu) = \eta_1$$
$$g_2(\sigma) = \eta_2$$
$$g_3(\nu) = \eta_3$$
$$g_4(\tau) = \eta_4.$$

The GAMLSS modelling framework is well-suited to modelling centiles for childhood growth data using parametric and semi-parametric methods. Stasinopoulos *et al*. (2017) state that centile curves should ideally be constructed using non-parametric smoothing methods in order to produce realistic and sensible curves. They caution that parametric methods such as polynomials and even fractional polynomials are often not flexible enough to adequately capture the attributes of the data. For smoothing methods, the degree of smoothing depends on smoothing parameters and the estimation of these smoothing parameters forms an important part of the modelling process.

Rigby and Stasinopoulos (2010) illustrated how GAMLSS could be used to derive child growth centile curves that vary as a function of age. Given $X = x$, denoting age, the response Y is modelled according to a chosen distribution with probability density function $f_Y(y|\mu,\sigma,\nu,\tau)$, where the parameters are modelled as smooth non-parametric functions of x, more specifically cubic splines, $h_k(x)$.

$$g_1(\mu) = h_1(x)$$
$$g_2(\sigma) = h_2(x)$$
$$g_3(\nu) = h_3(x)$$
$$g_4(\tau) = h_4(x).$$

The choice of monotonic link functions should be based on the requirements of the study and distribution chosen for the modelling but in general, the following link functions are conventionally used:

- $g_1(\mu)$ – identity link to allow it to have an additive effect and facilitate easy interpretation
- $g_2(\sigma)$ – log link such that the scale parameter will be non-negative
- $g_3(\nu)$ and $g_4(\tau)$ – identity and log link functions respectively.

A link function is chosen for each model parameter and each one should be appropriate for the model parameter's range (Stasinopoulos and Rigby, 2007).

The GAMLSS framework can accommodate a wide variety of distributions, including:

- Normal
- Box-Cox Cole and Green / LMS method
- Box-Cox Power exponential
- Box-Cox T

3.3.4 The LMS Method

The LMS method by Cole and Green (1992) is a special case of the GAMLSS which may be used when there is skewness in the data and there is no need to also model kurtosis. It assumes that the response has an underlying skew normal distribution and that at each value of the covariate, a Box-Cox power transformation can be applied to make the data normally distributed. The formula below transforms the response variable Y which results in the transformed variable Z with a standard normal distribution.

$$Z = \begin{cases} \frac{1}{\sigma\nu}\left[\left(\frac{y}{\mu}\right)^{\nu} - 1\right] & if\ \nu \neq 0 \\ \frac{1}{\sigma}\ log\left(\frac{y}{\mu}\right) & if\ \nu = 0 \end{cases}$$

where $\mu > 0, \sigma > 0, -\infty < \nu < \infty$
(Cole and Green, 1992).

When using the LMS method to estimate centile curves, we can allow the parameters to change as smooth functions of time and the parameters ν, μ, and σ in the formula above may be replaced by the smooth curves *L(t)*, *M(t)* and *S(t)*. At each value of the covariate, the distribution of the response is summarised by these three smooth curves, representing the Box-Cox power, the median and the coefficient of variation respectively. The formula may be written as:

$$z = \frac{[y/M(t)]^{L(t)} - 1}{L(t)S(t)} \qquad when\ L(t)\ \neq\ 0$$

$$z = \frac{\log[y/M(t)] - 1}{S(t)} \qquad when\ L(t)\ =\ 0$$

(Cole and Green, 1992).

The *L*, *M* and *S* curves are natural cubic splines which are estimated using maximum penalised likelihood. For *n* independent measurements y_i at covariate values t_i, the log likelihood is given by

$$l = l(L, M, S) = \sum_{i=1}^{n} (L(t_i) \log \frac{y_i}{M(t_i)} - \log S(t_i) - \frac{1}{2} z_i^2)$$

and the penalised likelihood is given by

$$l - \frac{1}{2}\alpha_\lambda \int \{L''(t)\}^2 dt \; - \frac{1}{2}\alpha_\mu \int \{M''(t)\}^2 dt \; - \frac{1}{2}\alpha_\sigma \int \{S(t)\}^2 dt \,,$$

where α_λ, α_μ and α_σ are the smoothing parameters which need to be chosen to fit the model. Alternatively, the equivalent degrees of freedom can be specified.

After fitting the model, centiles may be calculated by rearranging the z-score formula to give

$$C_{100\alpha}(t) = M(t)[1 + L(t)S(t)Z_\alpha]^{1/L(t)} \qquad L(t) \neq 0$$
$$C_{100\alpha}(t) = M(t)\exp[S(t)Z_\alpha] \qquad L(t) = 0$$

where Z_α is the normal equivalent deviate of size α
(Cole and Green, 1992).

3.3.5 The Box-Cox Power Exponential Method

Rigby and Stasinopoulos (2004) developed the Box-Cox Power Exponential method to provide a way to model both skewness and kurtosis in growth data. The model is denoted by *BCPE(μ, σ, ν, τ)* where *μ, σ, ν, τ* are the location (median), scale (coefficient of variation), skewness and kurtosis parameters respectively.

A variable *Y* with a BCPE distribution may be represented through the transformed variable *Z*, given by

$$Z = \begin{cases} \frac{1}{\sigma\nu}\left[\left(\frac{Y}{\mu}\right)^{\nu} - 1\right], & \nu \neq 0 \\ \frac{1}{\sigma} \, log\left(\frac{Y}{\mu}\right), & \nu = 0 \end{cases}$$

for $0 < Y < \infty, \mu > 0, \sigma > 0$ and where Z is a random variable with a standard power exponential distribution with power parameter $\tau > 0$. The probability density function of Z is given by

$$f_Z(z) = \frac{\tau}{c\, 2^{\left(1+\frac{1}{\tau}\right)}\Gamma(\frac{1}{\tau})}\, exp\left\{-0.5\left|\frac{Z}{c}\right|^{\tau}\right\}$$

$$-\infty < Z < \infty, \tau > 0$$

$$c^2 = 2^{-2/\tau}\Gamma\left(1/\tau\right)[\Gamma(3/\tau)]^{-1}$$

(Rigby and Stasinopoulos, 2004).

The benefit of this parameterisation is that Z has a mean and standard deviation of 0 and 1 respectively for all $\tau > 0$. When $\tau = 2$ then $f_Z(z)$ is the standard normal distribution so the LMS method is a special case of the BCPE method with a power parameter of 2.

We can then write the probability density function of Y, a $BCPE(\mu, \sigma, \nu, \tau)$ random variable, as

$$f_Y(y) = f_Z(z)\left|\frac{dz}{dy}\right| = \frac{y^{\nu-1}}{\mu^{\nu}\sigma}\, f_Z(z).$$

As with the LMS method, the centiles may be calculated using a simple formula:

$$C_{100\alpha}(t) = M(t)[1 + L(t)S(t)Z_\alpha]^{1/L(t)} \qquad \text{if } L(t) \neq 0$$

$$C_{100\alpha}(t) = M(t)\exp\left[S(t)Z_\alpha\right] \qquad \text{if } L(t) = 0$$

In this case, however, $Z_\alpha = F_Z^{-1}(\alpha)$ is the 100α centile of a standard power exponential distribution with power parameter τ.

$$Z_\alpha = \begin{cases} -c[2F_S^{-1}(1-2\alpha)]^{\frac{1}{\tau}} & for\ \alpha \leq 0.5 \\ c[2F_S^{-1}(2\alpha-1)]^{\frac{1}{\tau}} & for\ \alpha > 0.5 \end{cases}$$

where F_S^{-1} is the inverse cumulative distribution function of S, where S has a gamma distribution with probability density function given by

$$f_S(s) = \frac{s^{\frac{1}{\tau}-1}\exp(-s)}{\Gamma(\frac{1}{\tau})}$$

(Rigby and Stasinopoulos, 2004).

Rigby and Stasinopoulos (2004) applied the BCPE method to model the body mass index data from the Fourth Dutch Growth Study which is a cross-sectional study of the growth and development of the Dutch population between 0 and 21 years of age. This data had previously been analysed by van Buuren and Fredriks (2001) who found that their models for weight and head circumference exhibited clear evidence of kurtosis in the residuals. The BCPE method resulted in a much better fit to the data.

3.3.6 The Box-Cox t Method

The Box-Cox t (BCT) distribution was introduced by Rigby and Stasinopoulos (2006) for modelling a response variable which exhibits both skewness and leptokurtosis. The four parameters of the distribution, denoted $BCT(\mu, \sigma, \nu, \tau)$, represent the location, scale, skewness and kurtosis respectively.

For a random variable *Y*, the Box-Cox t distribution is defined through the transformed variable *Z*, given by

$$Z = \begin{cases} \frac{1}{\sigma\nu}\left[\left(\frac{Y}{\mu}\right)^{\nu} - 1\right] & \nu \neq 0 \\ \frac{1}{\sigma} log\left(\frac{Y}{\mu}\right) & \nu = 0 \end{cases}$$

for $0 < Y < \infty, \mu > 0, \sigma > 0$
(Rigby and Stasinopoulos, 2006).

It is assumed that the random variable *Z* follows a truncated *t* distribution with τ degrees of freedom with $\tau > 0$. The probability density function of $Y \sim BCT(\mu, \sigma, \nu, \tau)$ is given by

$$f_Y(y) = f_Z(z)\left|\frac{dz}{dy}\right| = \frac{y^{\nu-1}}{\mu^{\nu}\sigma} f_Z(z)$$

where $f_Z(z)$ is the probability density function of the transformed variable *Z* given by

$$f_Z(z) = \frac{f_T(z)}{F_T(1/\sigma|\nu|)}$$

for

$-1/(\sigma\nu) < Z < \infty$ if $\nu > 0$,

$-\infty < Z < -1/(\sigma\nu)$ if $\nu < 0$,

and where $\sigma > 0$ and $-\infty < \nu < \infty$

(Rigby and Stasinopoulos, 2006).

The 100α centiles of the random variable *Y* may be calculated exactly using the formula

$$C_{100\alpha} = \begin{cases} \mu[1 + \sigma\nu z_\alpha]^{\frac{1}{\nu}} & \nu \neq 0 \\ \mu \, exp[\sigma z_\alpha] & \nu = 0 \end{cases}$$

$$z_\alpha = \begin{cases} F_T^{-1}[\alpha F_T(1/\sigma|\nu|)] & \nu \leq 0 \\ F_T^{-1}[1 - (1-\alpha)F_T(1/\sigma|\nu|)] & \nu > 0 \end{cases}$$

where F_T^{-1} is the inverse cumulative distribution function of $T \sim t_\tau$

(Rigby and Stasinopoulos, 2006).

3.3.7 Model selection

Non-nested GAMLSS models may be compared using the Generalised Akaike Information Criterion (GAIC) (Stasinopoulos *et al*., 2017). This criterion penalises overfitting by penalising each effective degree of freedom with a fixed penalty # and is given by:

$$GAIC(\#) = GD + \# \, df$$
$$= -2l(\hat{\theta}) + \# \, df$$

where GD is the fitted global deviance. The preferred model is the one with the smallest GAIC(#).

Using penalties of *#=2* and *#=log(n)* equate to using the Akaike Information Criterion (AIC) (Akaike, 1974) and the Schwarz Bayesian Criterion/Bayesian Information Criterion (BIC) (Schwarz, 1978) respectively as these are special cases of the GAIC. Using the GAIC offers flexibility in the choice of penalty which may be adjusted based on the modelling problem. It allows the researcher to test the sensitivity of the model to different penalties. The higher the penalty, the lower the chosen degrees of freedom and therefore the smoother the fitted centile curves.

Van Buuren and Fredriks (2001) provide a breakdown of useful diagnostic tools for assessing the suitability of growth reference curves.

- **Visually assess the shape of the curves**
  An experienced researcher may be able to determine the suitability of the curves based on expectations about what the shape should look like.
- **Overlaying derived reference curves on observed data**
  While not helpful for selecting from a pool of potential models, it may aid in identifying outliers as well as obvious deficiencies in the model.
- **Overlaying empirical centiles on derived centiles**
  To get the empirical centiles, the age range is divided into groups and the empirical centiles are computed for each age group and then plotted to form a curve. There should be close agreement between the empirical and derived curves. Van Buuren and Fredriks (2001) urged that caution be taken when using the raw as opposed to standardised data because for data exhibiting non-constant variance across age, the same distance will have different meanings at different ages.
- **Observed and expected counts**
  The observed proportions of observations lying below each centile can be compared to the expected proportions based on the fitted model.
- **Statistical tests**
  These may accompany parametric modelling to check that, after transformation, the data does in fact follow the assumed distribution. For example, the LMS method assumes that the transformed variable Z follows a standard normal distribution at all ages. Tests such as the Shapiro-Wilk W test and Kolmogorov-Smirnov tests may be used to confirm these assumptions.
- **Quantile-Quantile (Q-Q) plots of the z-scores**
  The Q-Q plot plots the empirical quantiles against the theoretical ones in order to highlight deviations between the two.
- **Worm plot**

The worm plot was developed by van Buuren and Fredriks (2001) to serve as a residual diagnostic tool to assess model fit and thereby guide the model selection process in deriving age-related growth reference curves.

A worm plot consists of a series of detrended Q-Q plots, each one drawn for a different level of a covariate such as age. Van Buuren and Fredriks (2001) explain that the position of an individual observation on the vertical axis represents the difference between its location in the theoretical and empirical distributions and when plotted all together, the points form a worm-like string. Flat worms are an indication that the model fits the data well and deviations from the horizontal shape alert the user to locations at which the model fit is lacking. Table 3.1 is provided to aid interpretation of various patterns which may be observed in the worm plot.

Table 3.1: Interpretation of patterns observed in a worm plot.

Shape	**Moment**	**If the**	**Then the**
Intercept	Mean	worm passes above the origin	fitted mean is too small
		worm passes below the origin	fitted mean is too large
Slope	Variance	worm has a positive slope	fitted variance is too small
		worm has a negative slope	fitted variance is too large
Parabola	Skewness	worm has a U-shape	fitted distribution is too skew to the left
		worm has an inverted U-shape	fitted distribution is too skew to the right
S-curve	Kurtosis	worm has an S-shape on the left bent down	tails of the fitted distribution are too light
		worm has an S-shape on the left bent up	tails of the fitted distribution are too heavy

(van Buuren and Fredriks, 2001).

Van Buuren (2007) discusses two potential causes of inadequate model fit for a parametric model such as the LMS model.

1. Choosing the incorrect conditional distribution
2. Biased estimates for the mean, variance, skewness and kurtosis parameters as a result of over-smoothing.

   The model fitting process for parametric models involves determining the best values for the effective degrees of freedom of the fitted curves. A low degrees of freedom value sacrifices model fit in the interest of smoothness of the curve at the risk of smoothing out important growth patterns. A high degrees of freedom value ensures that the model follows the data closely but the result may be wiggly growth reference

curves as a consequence of picking up the noise in the data rather than just the underlying growth trends.

The worm plot is an apt tool for aiding the decision of how many degrees of freedom to use, with the user increasing the degrees of freedom until the further improvement in the worm plot is not meaningful.

3.4 Non-Parametric Methods

3.4.1 Quantile regression

Quantile regression does not make any assumptions about the distribution of the data and therefore serves as an alternative to parametric modelling. Since it does not impose any global constraints, it is capable of modelling data which exhibits different behaviour across different age intervals.

Koenker and Bassett (1978) formulated quantile regression by building on from the optimisation interpretation of ordinary sample quantiles in which the function

$$\sigma(\beta) = \sum_{i=1}^{n} \rho_\tau(y_i - \beta)$$

$$\rho_\tau(u) = u\big(\tau - I(u < 0)\big)$$

is minimised over the scalar β to yield the τth sample quantile $\hat{\beta} = F_y^{-1}(\tau)$ of $y = \{y_1, y_2, \dots, y_n\}$.

In quantile regression, β is replaced by a linear function of the covariates x which produces the objective function

$$\sigma(\beta) = \sum_{i=1}^{n} \rho_\tau(y_i - x_i'\beta)$$

$$\rho_\tau(u) = u\big(\tau - I(u < 0)\big)$$

$$= \begin{cases} \tau\, u & u \geq 0 \\ (\tau - 1)u & u < 0 \end{cases}$$

which, when minimised, yields the estimate $x_i'\hat{\beta}$, described by van Buuren (2007) as the τth conditional quantile function of *y* given *x*. The function ρ_τ is a piecewise linear function which ensures that the number of observations above and below the fitted curve remains balanced.

Wei *et al*. (2004) explain that the objective function may be represented more generally as

$$\sum_{i=1}^{n} \rho_\tau(y_i - g(x_i))$$

Where g belongs to a family of smooth functions G. They caution that children's heights are known to deviate from normality and growth is non-linear, while other measures such as weight may potentially be even more problematic. Considering this, cubic splines are a good choice for the family of functions G. For a chosen set of knots, a cubic spline may be written as a linear combination of basis splines $h_j, j = 1, 2, \dots, J$. That is,

$$g(x_i) = \sum_{j=1}^{J} \beta_j h_j(x_i)$$

In this case, quantile regression involves estimating the $\beta = (\beta_1, \beta_2, \dots, \beta_J)^T$ which minimises the objective function

$$\sum_{i=1}^{n} \rho_\tau(y_i - \sum_{j=1}^{J} \beta_j h_j(x_i))$$

Wei *et al*. (2006) state that this problem may be solved quickly and easily by linear programming.

3.4.2 Model selection

Since quantile regression is distribution agnostic, incorrectly specifying the conditional distribution of the data is not a potential cause of poor model fit. The risk of bias remains, however, if the component $x_i'\beta$ is misspecified (Van Buuren, 2007).

The quantile regression modelling process involves the selection of a set of knots for which the basis splines are constructed. Wei *et al*. (2006) recommends using fairly evenly spaced knots with extra knots during periods of rapid growth such as in early infancy, and to keep the set of knots the same for different centiles. Advice in the literature regarding the selection of knots is scarce. Van Buuren (2007) proposed using the worm plot to diagnose the fit of and compare competing quantile regression models and demonstrated their use using growth data from the Fourth Dutch Growth Study.

As discussed above, the conventional worm plot graphically depicts discrepancies between the empirical and theoretical distributions. Using worm plots to assess the fit of quantile regression models presents a challenge because quantile regression does not assume any underlying distribution from which the theoretical quantities are drawn.

Van Buuren (2007) therefore proposed a distribution free version of the worm plot as a diagnostic tool for quantile regression. This is based on the idea that if the quantile regression model fits the data then the empirical quantiles must be uniformly distributed. The worm plot is used to establish if this uniformity holds and thereby indicates whether the model fits the data well.

Van Buuren (2007) details the steps involved in calculating the points required to produce a worm plot for quantile regression. Essentially, the fitted values from quantile regression models fitted for various values of τ are used to construct an approximate numerically calculated CDF which is then used to calculate the normalised quantile residuals which are used to produce the worm plot.

Quantile regression is increasing in popularity as a method for deriving growth reference curves with a number of authors supporting its use: Gannoun *et al*. (2002), He (1997), Schnabel and Eilers (2012), Wei *et al*. (2006) and Muggeo (2021).

As discussed by van Buuren (2007), quantile regression does however have some disadvantages, the main one being that there is no explicit formula for calculating a quantile or z-score for an individual measurement. This was an important requirement for the WHO when considering potential approaches for constructing growth reference curves (Borghi *et al.*, 2006) and so quantile regression was disregarded on this basis.

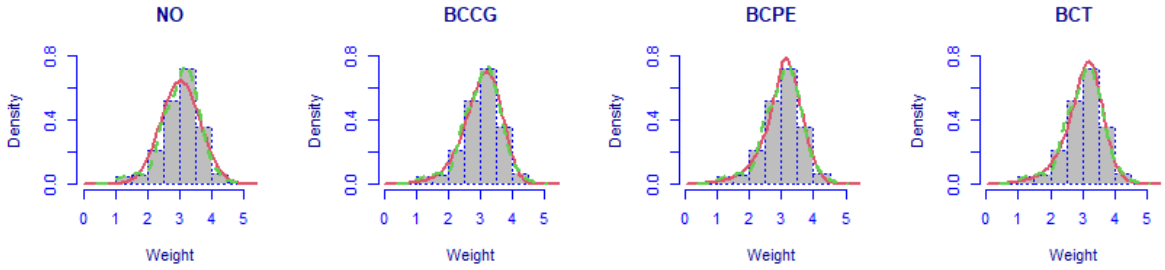
Two further problems experienced by users of quantile regression are the potential for irregular curves near the extremes of the data, as well as the tendency for quantile curves to cross, making the curves difficult to use and interpret. Van Buuren (2007) proposed using different levels of smoothing for different levels of τ to combat the issue of irregular curves. Schnabel and Eilers (2013) proposed the use of quantile sheets regression to combat both the irregular curves at the extremes and crossing curves. Their approach involves simultaneous estimation of the quantiles and the use of smoothing parameters in the direction of the explanatory variable as well as the response variable.

4 Derivation of growth curves

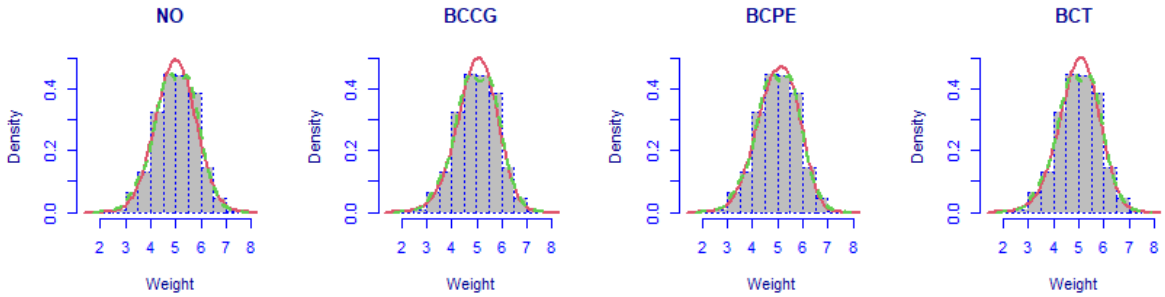
4.1 Weight-for-age: boys

Theoretical densities of the Normal, BCCG, BCPE and BCT distribution were plotted along with the observed density histograms of weight for boys (Figure 4.1) in order to ascertain which distributions will provide a good fit to the data. The BCCG, BCPE and BCT distributions appear to fit the data much better than the normal distribution which suggests the need for modelling skewness and kurtosis.

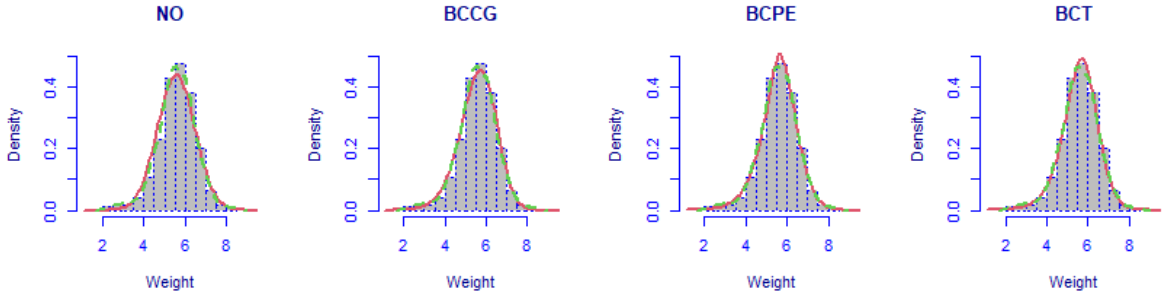
(a) Birth



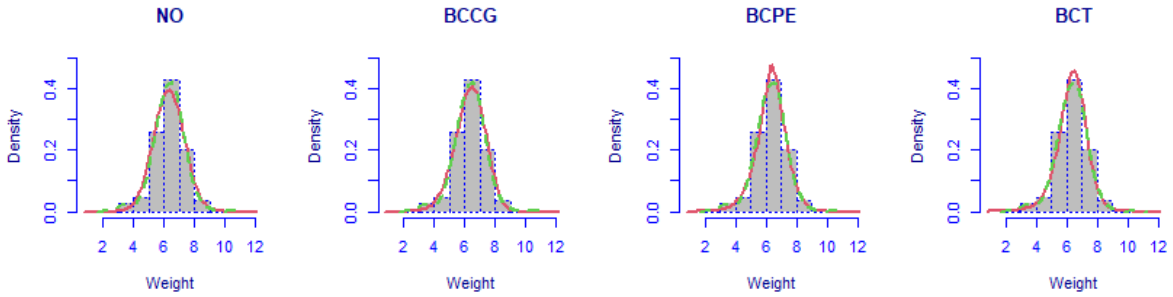
(b) 6-10 weeks



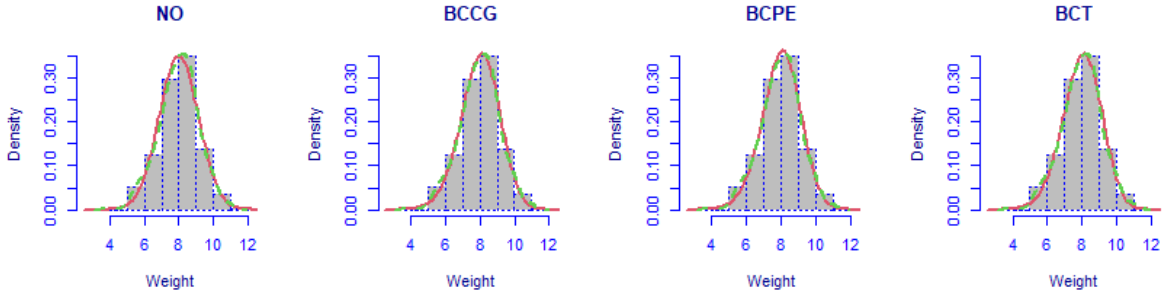
(c) 10 weeks



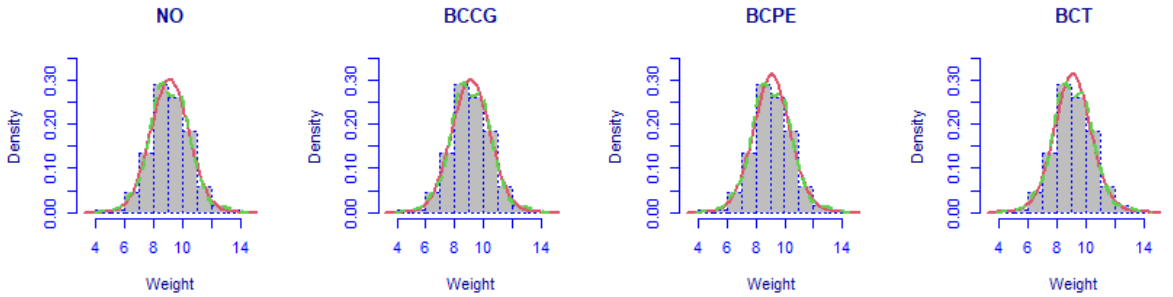
(d) 14 weeks



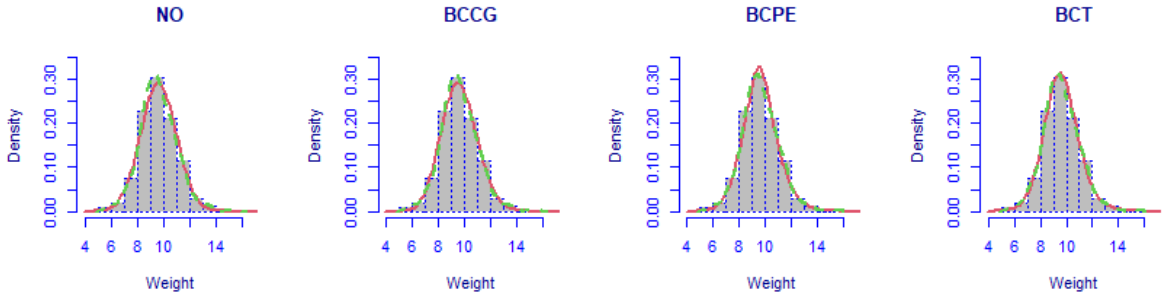
(e) 6 months



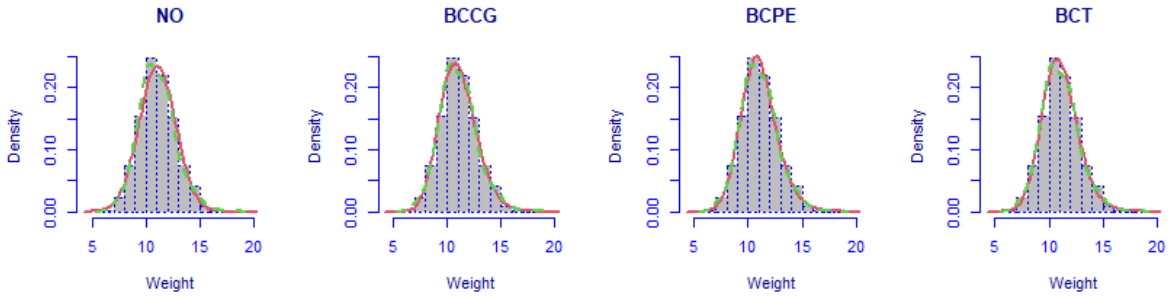
(f) 9 months



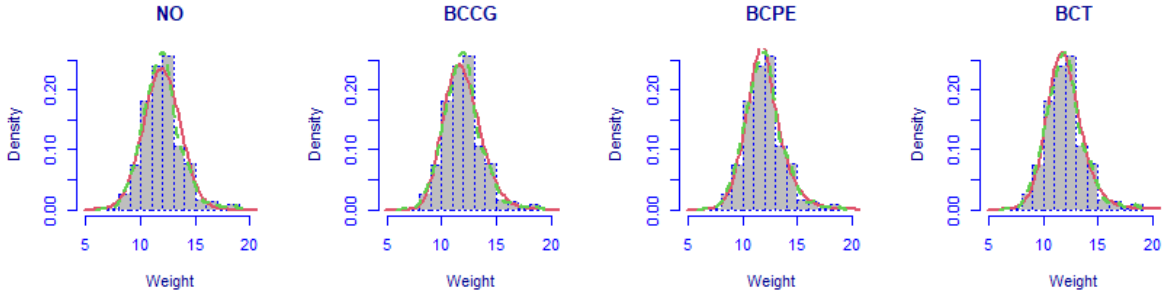
(g) 12 months



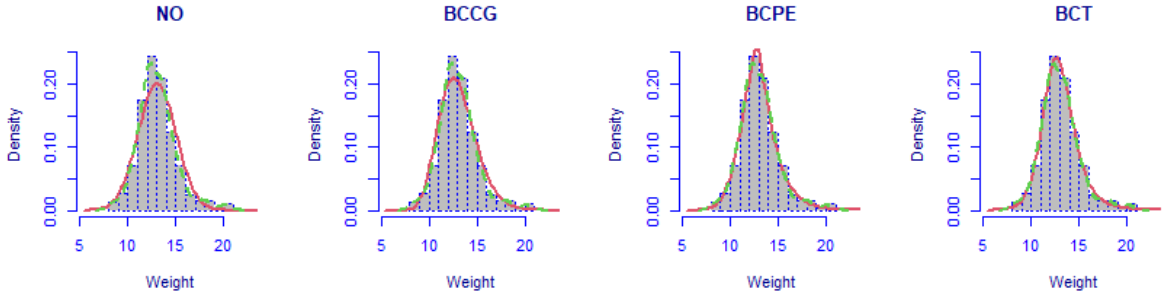
(h) 18 months



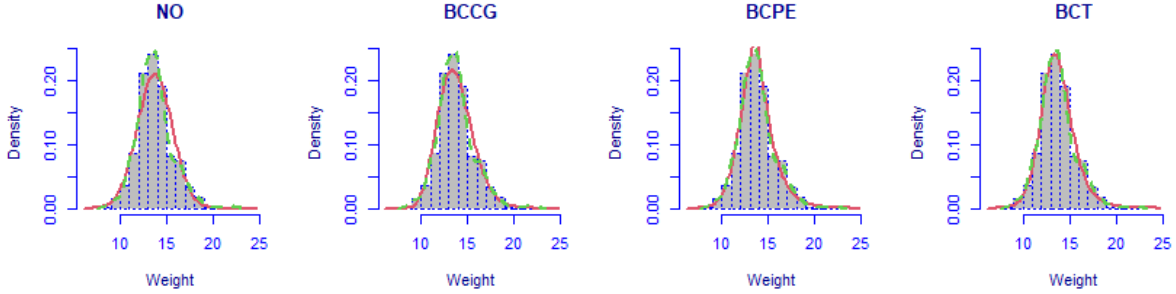
(i) 24 months



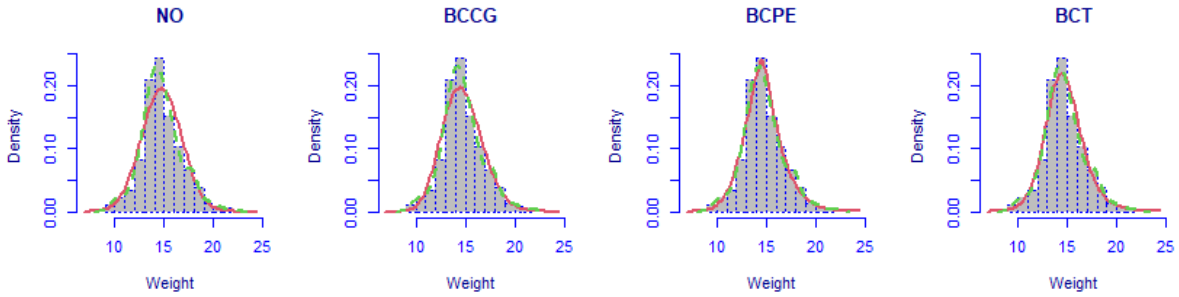
(j) 30 months



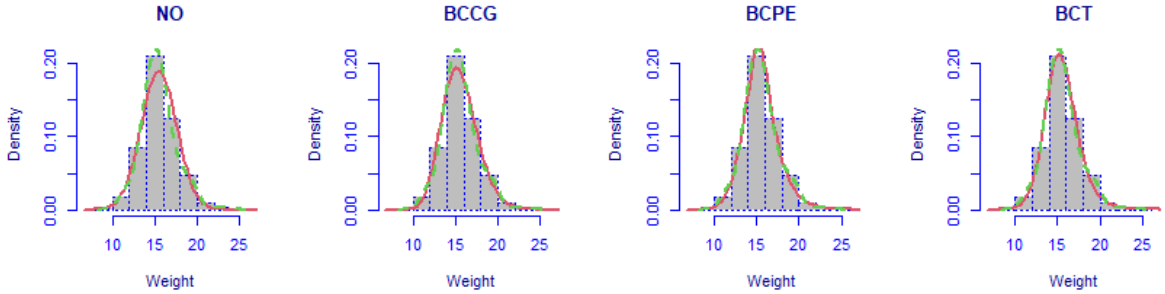
(k) 36 months



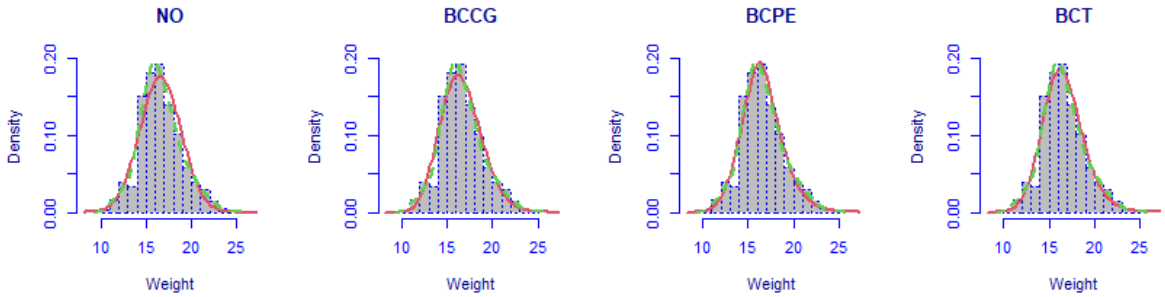
(l) 42 months



(m) 48 months



(n) 54 months



(o) 60 months

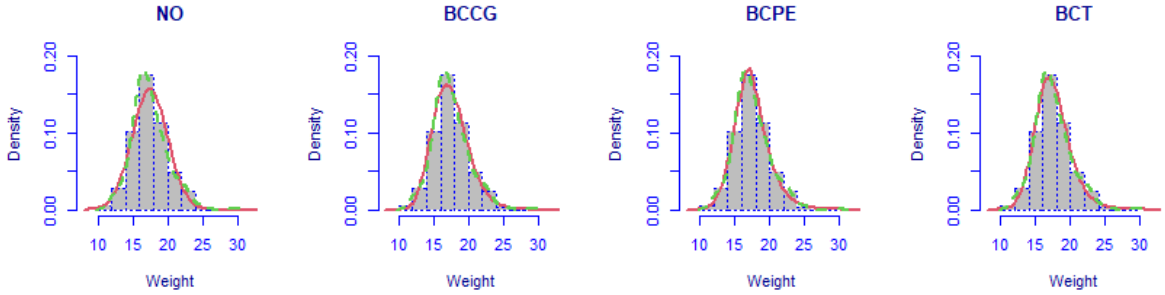


Figure 4.1: Density histograms of weight for boys with theoretical densities of normal, BCCG, BCPE, BCT distributions overlaid.

*Green dotted line is a non-parametric kernel density
*Red line is the theoretical density

4.1.1 GAMLSS models with curve smoothing by fractional polynomials

The conventional approach to fractional polynomial regression (described in section 3.3.1) assumes that the data follows a normal distribution at all ages. The first step to fitting a fractional polynomial model involves choosing a suitable mean model, followed by a separate analysis to model the variability around the mean. If the residuals are normally distributed then the absolute residuals follow a half-normal distribution and the age specific standard deviation can be estimated by regressing the absolute residuals on age and then multiplying the fitted values by $\sqrt{pi/2}$. Since the distribution of boys' weights was shown to be non-normal, this approach cannot be used as is.

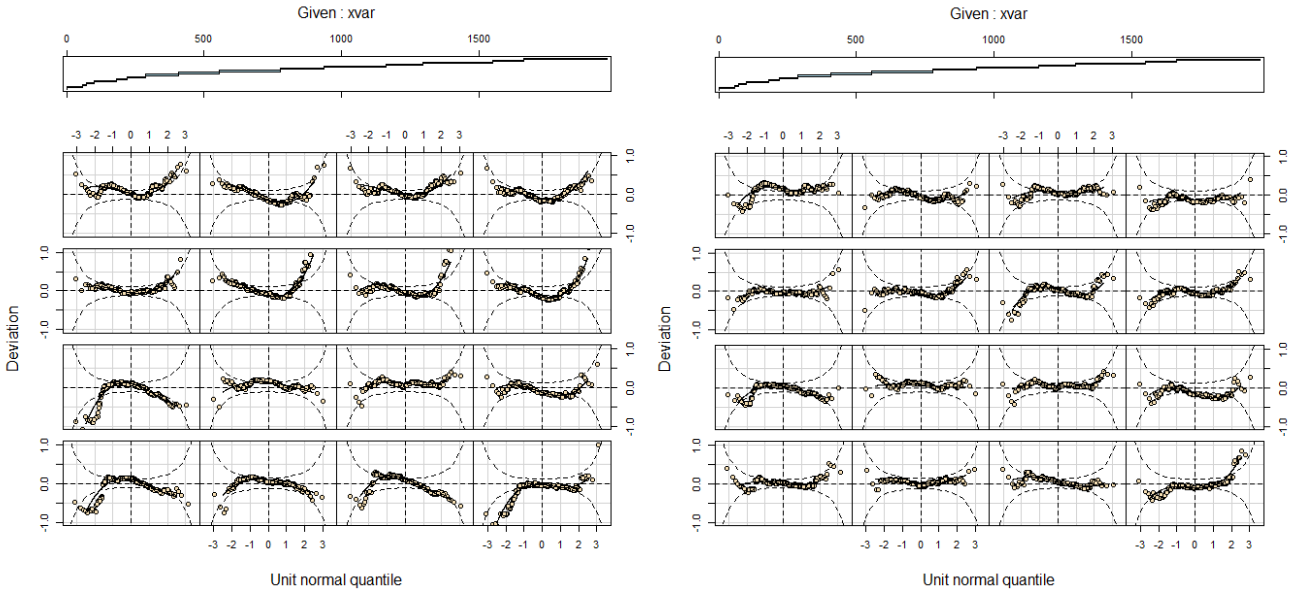
However, within the GAMLSS framework, all of the distribution parameters μ, σ, ν, τ may be modelled as linear and/or nonlinear parametric and/or nonparametric smoothing functions of the explanatory variables. Fractional polynomials may be fitted within this framework as additive "smoother" terms resulting in parametric models.

Fractional polynomials were used to model the relationship between the distribution parameters and age for each of the distributions: Normal (NO), Box-Cox Cole and Green (BCCG), BCPE, and BCT. Models with two and three fractional polynomial terms were investigated. Prior to the modelling, since the growth measurements begin at birth, 1 day was added to each age value to ensure that the positivity requirement is met. For each distribution, the best-fitting model was selected based on the GAIC with a penalty of 3. The best-fitting fractional polynomials of degree 2 and 3 for each distribution and its parameters are presented in Table 4.1.

Table 4.1: Best fitting fractional polynomials for each distribution.

Distribution	Parameter	Model formula
Normal, 2 terms	μ	$\mu = \beta_0 + \beta_1 age^{0.5} + \beta_2 age$
	σ	$\sigma = \beta_0 + \beta_1 age^{-0.5} + \beta_2 \log(age)$
Normal, 3 terms	μ	$\mu = \beta_0 + \beta_1 age^{-0.5} + \beta_2 \log(age) + \beta_3 age^2$
	σ	$\sigma = \beta_0 + \beta_1 age^{-0.5} + \beta_2 age^{-0.5} \log(age) + \beta_3 age^2$
BCCG, 2 terms	μ	$\mu = \beta_0 + \beta_1 age^{-2} + \beta_2 \log(age)$
	σ	$\sigma = \beta_0 + \beta_1 age^{-2} + \beta_2 age^{0.5}$
	ν	$\nu = \beta_0 + \beta_1 age^{-1} + \beta_2 \log(age)$
BCCG, 3 terms	μ	$\mu = \beta_0 + \beta_1 age^{-0.5} + \beta_2 age^{-0.5} \log(age) + \beta_3 age^2$
	σ	$\sigma = \beta_0 + \beta_1 age^{-0.5} + \beta_2 age^2 + \beta_3 age^3$
	ν	$\nu = \beta_0 + \beta_1 age^2 + \beta_2 age^2 \log(age) + \beta_3 age^2 (\log age)^2$
BCPE, 2 terms	μ	$\mu = \beta_0 + \beta_1 \log(age) + \beta_2 (\log age)^2$
	σ	$\sigma = \beta_0 + \beta_1 \log(age) + \beta_2 age^3$
	ν	$\nu = \beta_0 + \beta_1 \log(age) + \beta_2 (\log age)^2$
	τ	$\tau = \beta_0 + \beta_1 age^3 + \beta_2 age^3 \log(age)$
BCPE, 3 terms	μ	$\mu = \beta_0 + \beta_1 age^{-0.5} + \beta_2 \log(age) + \beta_3 age^2$
	σ	$\sigma = \beta_0 + \beta_1 age^{-2} + \beta_2 age^{-2} * \log(age) + \beta_3 age^{0.5}$
	ν	$\nu = \beta_0 + \beta_1 age^2 + \beta_2 age^2 \log(age) + \beta_3 age^2 (\log age)^2$
	τ	$\tau = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 age^2 \log(age)$
BCT, 2 terms	μ	$\mu = \beta_0 + \beta_1 \log(age) + \beta_2 (\log age)^2$
	σ	$\sigma = \beta_0 + \beta_1 \log(age) + \beta_2 age^3$
	ν	$\nu = \beta_0 + \beta_1 \log(age) + \beta_2 (\log age)^2$
	τ	$\tau = \beta_0 + \beta_1 age^3 + \beta_2 age^3 \log(age)$
BCT, 3 terms	μ	$\mu = \beta_0 + \beta_1 age^{-0.5} + \beta_2 \log(age) + \beta_3 age^2$
	σ	$\sigma = \beta_0 + \beta_1 age^{-2} + \beta_2 age^{-2} * \log(age) + \beta_3 age^{0.5}$
	ν	$\nu = \beta_0 + \beta_1 age^2 + \beta_2 age^2 \log(age) + \beta_3 age^2 (\log age)^2$
	τ	$\tau = \beta_0 + \beta_1 age^3 + \beta_2 age^3 \log(age) + \beta_3 age^3 (\log age)^2$

As expected, the Normal and BCCG distributions do not fit the data adequately as indicated by the worm plots in Figure 4.2. The worms for the normal distribution are particularly poor: most are U-shaped or inverse U-shaped which suggests a strong presence of residual skewness.



(a) (b)

Figure 4.2: Worm plots for (a) Normal and (b) BCCG distributions with curve smoothing by fractional polynomials of degree 3.

A comparison of the fractional polynomial models based on the AIC, GAIC(3) and BIC is displayed in Table 4.2. The best-fitting model was the BCPE distribution with smoothing by fractional polynomials with three terms. It produced the lowest AIC, GAIC(3) as well as BIC.

Table 4.2: Comparison of fractional polynomial models.

	Distribution	**No. of FP terms**	**AIC**	**GAIC(3)**	**BIC**
1	NO	2	25662.73	25672.73	25732.73
2	NO	3	25411.09	25425.09	25509.09
3	BCCG	2	25313.36	25328.36	25418.36
4	BCCG	3	25177.75	25198.75	25324.75
5	BCPE	2	25366.29	25386.29	25506.29
6	**BCPE**	**3**	**25094.64**	**25122.64**	**25290.64**
7	BCT	2	25373.30	25393.30	25513.30
8	BCT	3	25103.38	25131.38	25299.38

Residual diagnostics did not reveal any significant problems with the fit. Figure 4.3 shows a random scatter in the plot of quantile residuals vs fitted values and the Normal Q-Q plot shows no significant deviations. The worm plots for the BCPE distribution (Figure 4.4) are

relatively flat and indicate improved fit over the Normal and BCCG distributions. There are a few points which lie outside of the 95% confidence bands in a few of the age intervals but not more than is to be expected.

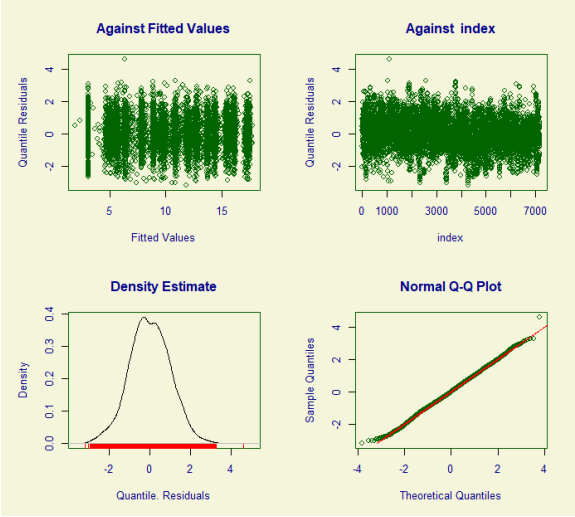


Figure 4.3: Residual diagnostics for BCPE model with fractional polynomial smoothing.

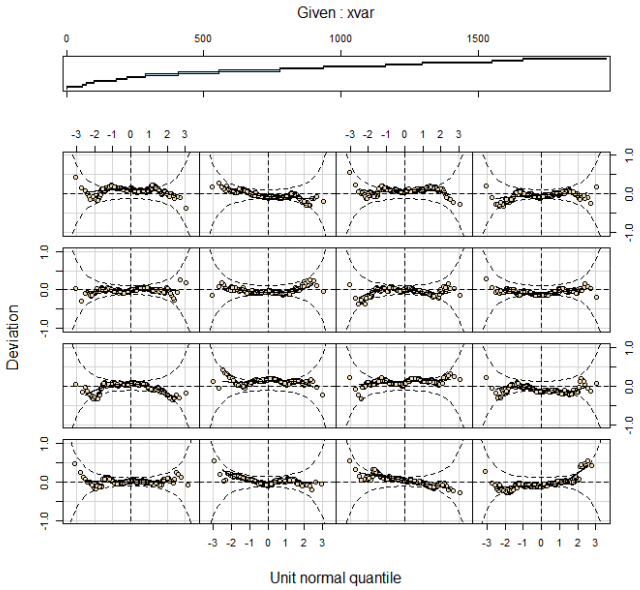


Figure 4.4: Worm plot for BCPE model with fractional polynomial smoothing.

Q-statistics were calculated to assess normality of residuals within age intervals. These are represented visually in Figure 4.5. Z values greater than 2 or less than -2 suggest significant inadequacies in the model. These correspond to the red and blue circles with squares in them. If greater than 2 then the residuals have a higher mean, variance, skewness or kurtosis than the standard normal distribution. Similarly, if less than 2 then the residuals have a lower

mean, variance, skewness or kurtosis than the standard normal distribution. The larger the circles, the more inadequate the model is. These problems may be addressed by increasing the number of degrees of freedom for the particular parameter. Analysis of the Q-statistics (or Z-statistics) in Figure 4.5 reveal a small number of age intervals in which the fit is not ideal however, overall the fit is considered to be adequate. The resulting centile curves shown in Figure 4.6 are not as smooth as desired.

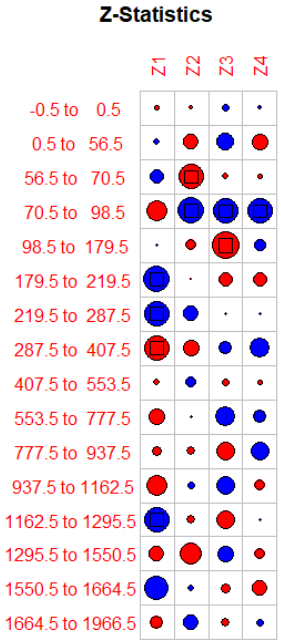


Figure 4.5: Visualising Q-statistics for BCPE model with fractional polynomial smoothing.

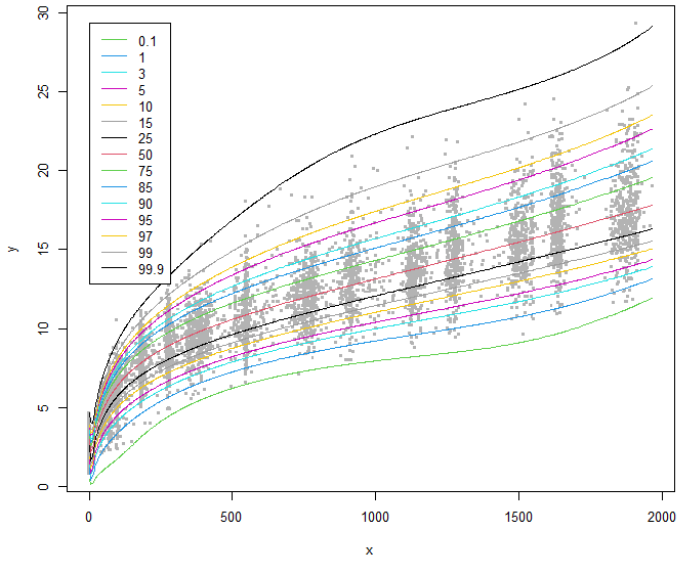


Figure 4.6: Centile curves for BCPE model with fractional polynomial smoothing.

4.1.2 GAMLSS models with Berkey-Reed mean model

The Berkey-Reed models make the assumption that the data are normally distributed with constant variance, an assumption which was shown to be violated by the children's weights. Figure 4.7 confirms this with evidence of poor model fit for the first order model. There is a fanning out of residuals in the plot of residuals versus fitted values indicating increasing variability with age and the Normal Q-Q plot shows obvious departures from normality.

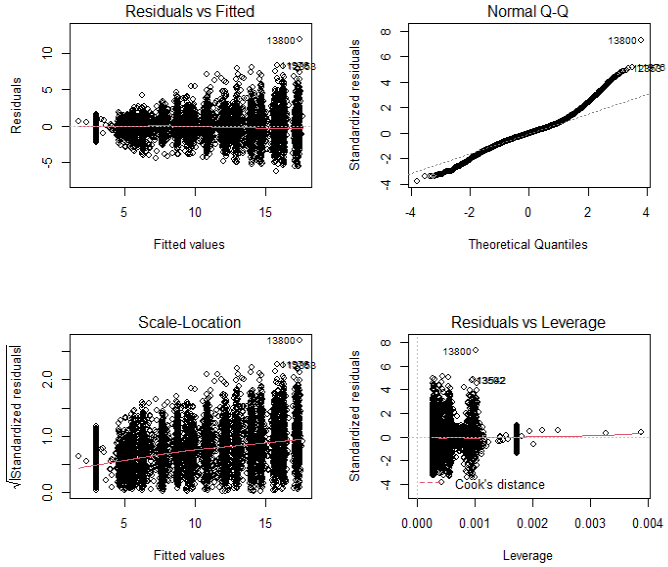
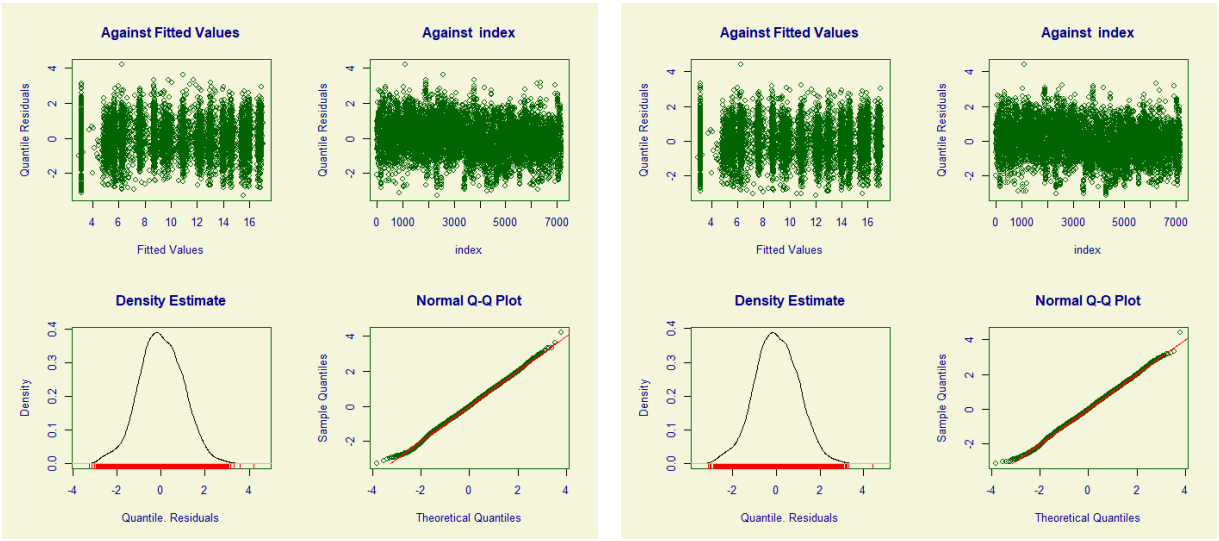


Figure 4.7: Residual diagnostics for Berkey-Reed First Order model assuming normality.

The approach taken was to use the Berkey-Reed models to model childhood growth within the GAMLSS framework such that the age-varying variance, skewness and kurtosis could be accounted for. The Reed models were used as the model for μ. The first approach was to allow the other three parameters to change linearly with age, resulting in parametric models and the second was to use cubic splines, resulting in semi-parametric models. Since the Berkey-Reed models are not defined at zero, 1 day was added to each age measurement.

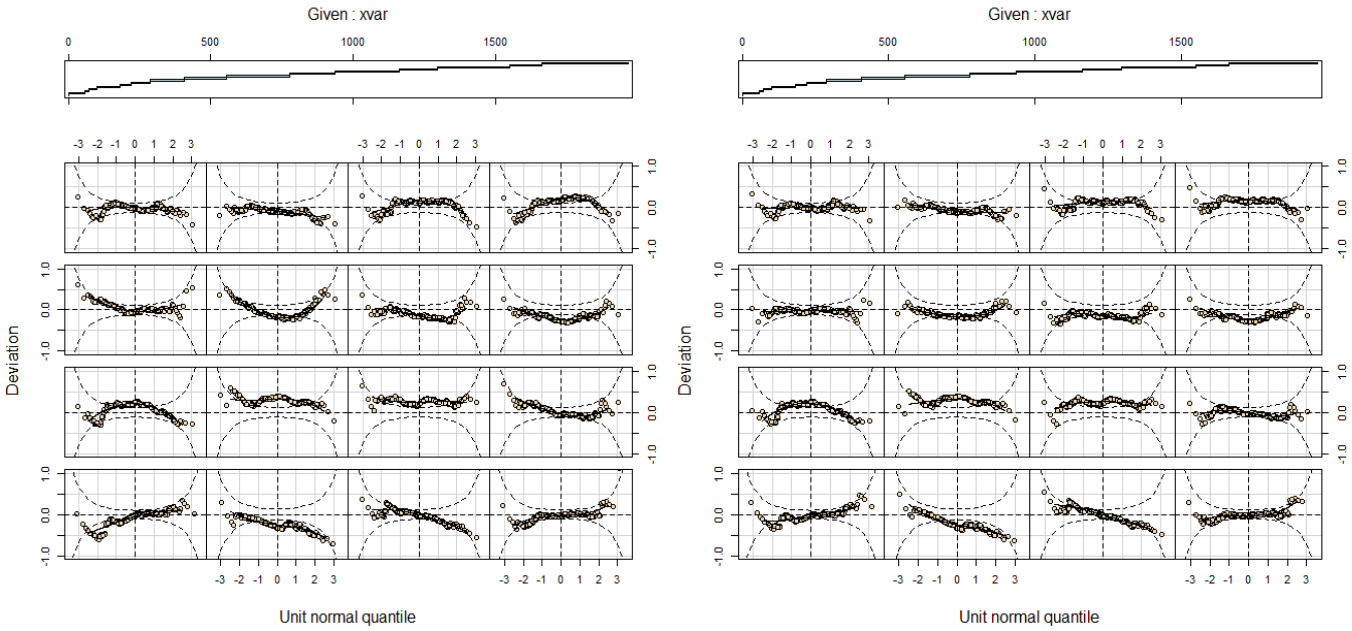
Figure 4.8 compares the residual diagnostics for the first order model where σ, ν and τ were linearly related to age (Figure 4.8 (a)) against the model which used cubic splines (Figure 4.8 (b)). The model fits are both a significant improvement over the traditional approach, however, as evidenced by the worm plots and Q-statistics in Figures 4.9 and 4.10, the skewness and kurtosis are better accounted for when modelling the relationship with age using cubic splines. It is also apparent from the worm plots and Q-statistics that the Berkey-

Reed first order model does not fit the μ parameter very well. A number of the worms are crossing above or below the origin indicating that the mean is being over or underestimated. The large absolute values of the Q-statistics confirm this.



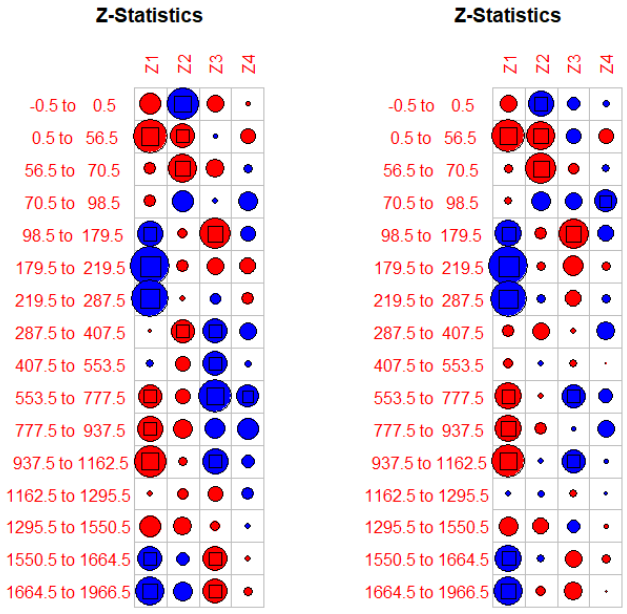
(a) **(b)**

Figure 4.8: Residual diagnostics for Berkey-Reed First Order models with σ, ν and τ linearly related to age (a) and modelled using cubic splines (b).



(a) **(b)**

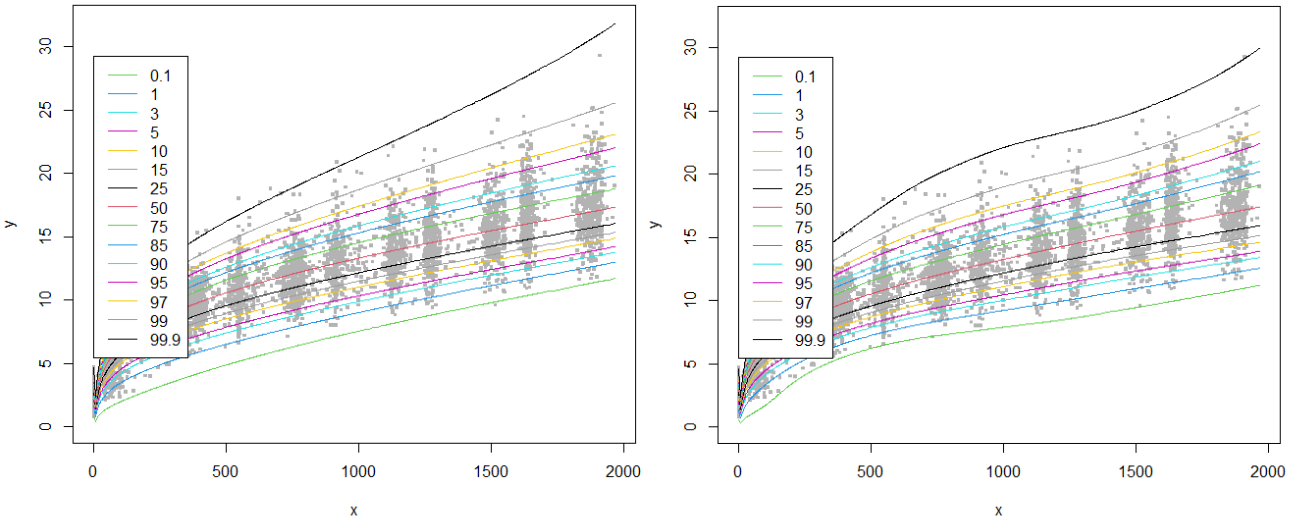
Figure 4.9: Worm plots for Berkey-Reed first order models with σ, ν and τ linearly related to age (a) and modelled using cubic splines (b).



(a) **(b)**

Figure 4.10: Q statistics for Berkey-Reed first order models with σ, ν and τ linearly related to age (a) and modelled using cubic splines (b).

The centile curves for the models are shown in Figure 6.11. Although the cubic splines model achieves better model fit, the centile curves are less smooth.



(a) **(b)**

Figure 4.11: Centile curves for Berkey-Reed first order models with σ, ν and τ linearly related to age (a) and modelled using cubic splines (b).

The same comparison was conducted for the Berkey-Reed second order model and no significant improvement in model fit was achieved over the first order model.

4.1.3 GAMLSS Models with cubic spline smoothing

Each parameter of a chosen distribution may be modelled as a cubic spline within the GAMLSS framework. Each cubic spline has a smoothing penalty associated with it which may be expressed in quadratic form. The trace of the corresponding matrix gives the effective degrees of freedom (edf) which is inversely related to the smoothing penalty.

Models are specified by the effective degrees of freedom of each parameter. The choice of effective degrees of freedom determines the degree of smoothing applied to each parameter. There is a trade-off between model fit and the degree of smoothness. To illustrate this, a Box-Cox Power Exponential model with low degrees of freedom, *BCPE(1, 1, 1, 1)* is compared to one with high degrees of freedom, *BCPE(20, 20, 20, 20)*.

The worm plots in Figure 4.12 show that the model with higher degrees of freedom fits the data much better – the worms are horizontal and pass through the origin. The low edf model performs poorly in terms of model fit. Many of the worms are U-shaped or S shaped indicating that the skewness and kurtosis have not been adequately modelled. Many of them cross the vertical line above or below the origin which is an indication that the mean parameter has also been inadequately modelled.

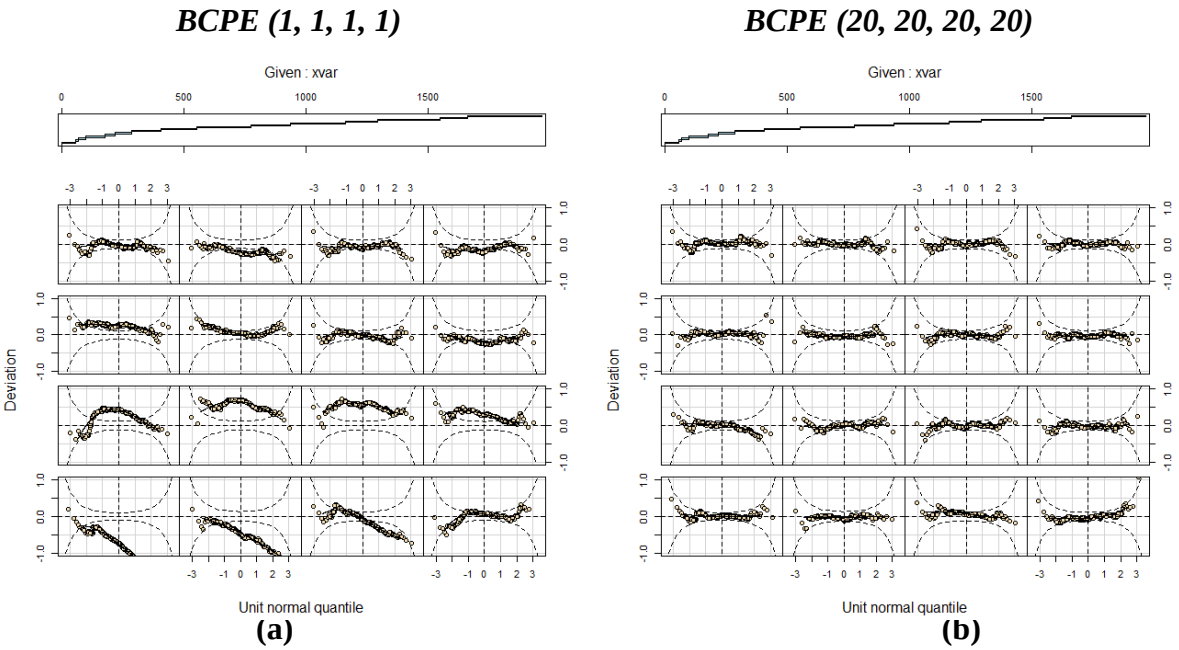
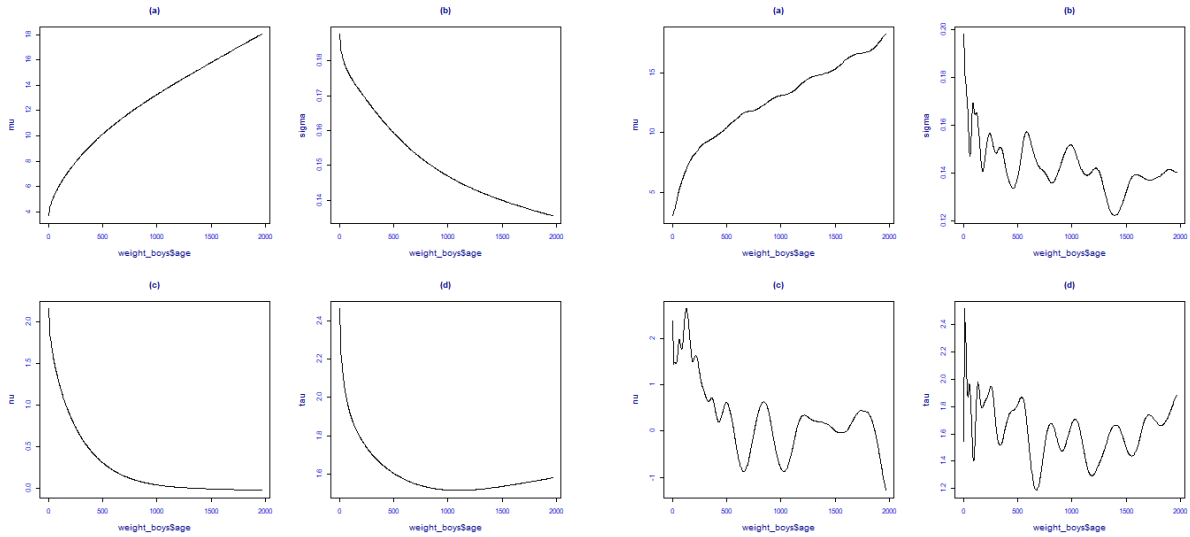


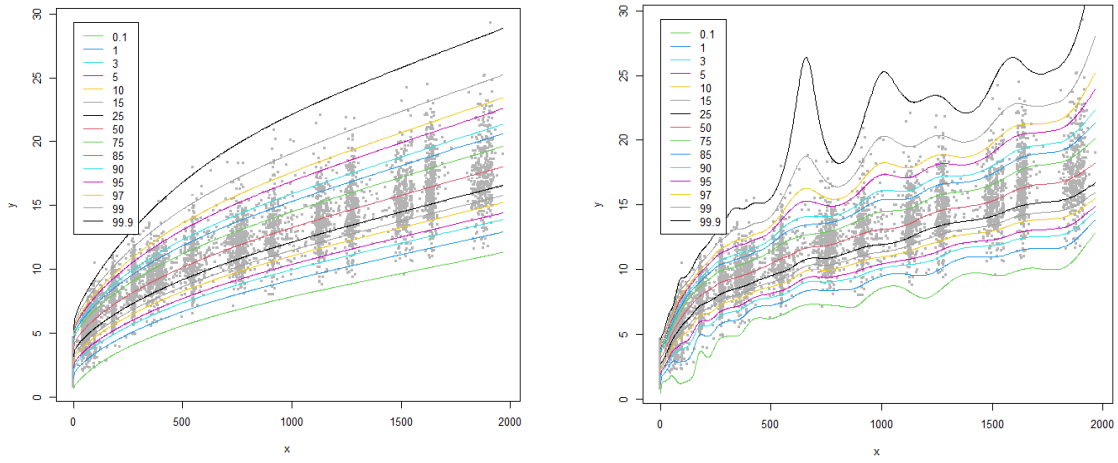
Figure 4.12: Worm plots for models (a) *BCPE(1, 1, 1, 1)* and (b) *BCPE(20, 20, 20, 20)*.

The corresponding fitted parameter curves in Figure 4.13 as well as the centiles in Figure 4.14 illustrate the impact of the degrees of freedom on the fitted models. The low edf model produces very smooth fitted curves and centile curves compared to the very wavy curves of the high edf model. It is clear that the low edf model is underfitting the data and smoothing out the biological growth pattern of the children too much. The high edf model shows significant overfitting with the curves picking up the noise in the data.



(a) **(b)**

Figure 4.13: Fitted parameter curves for (a) *BCPE(1, 1, 1, 1)* and (b) *BCPE(20, 20, 20, 20)*.



(a) **(b)**

Figure 4.14: Centile curves for models (a) *BCPE(1, 1, 1, 1)* and (b) *BCPE(20, 20, 20, 20)*.

Stasinopoulos *et al*. (2017) outline three approaches which may be used in combination to determine smoothing parameters:

- **Subjective but structured methods**

  Choose smoothing parameters based on prior knowledge and experience together with some broad guidelines.

- **Automatic methods**

  The selection of the smoothing parameters is done using an automatic procedure based on a chosen criterion.

- **Methods based on diagnostics**

  Smoothing parameters may be chosen such that the resultant model performs well according to some diagnostic tools. Examples are the worm plots developed by Van Buuren and Fredriks (2001) and Q statistics of Royston and Wright (2000).

4.1.3.1 Automatic selection of hyperparameters

Stasinopoulos *et al*. (2017) describe three automatic procedures for the selection of hyperparameters.

- **Method 1**

  Minimise the GAIC(k) over the five hyperparameters where k is the chosen fixed penalty. This is a global minimisation procedure. Rigby and Stasinopoulos (2006) used this method to select the hyperparameters when deriving centile curves for head circumferences of males in the Fourth Dutch Growth study. They compared various distributions as well as different values for the penalty k.

- **Method 2**

  Rigby and Stasinopoulos (2007) proposed a method which involves minimising the validation global deviance over the five hyperparameters. They found that this method resulted in overfitting of the dataset used. The steps are as follows:

  (1) Split data randomly into a training and testing set.

  (2) For a particular set of hyperparameters, fit the model to the training data. Calculate the validation global deviance, minus twice the log likelihood of the validation data.

  (3) Minimise the validation global deviance over all the sets of hyperparameters. This may be achieved using the R function **`optim()`** (R Core Team, 2021).

- **Method 3**

  The third method is a two-step procedure. In the first step, the appropriate transformation of the *x* variable is determined by assuming a simple normal distribution model with $g(\mu) = S_1(x^{\xi})^{\nu}$ and a constant σ and minimising the GAIC(*k*) over ξ for fixed penalty *k*. In the second step, models are fitted for various distributions and the parameters are estimated using a local maximum likelihood procedure. The benefit of this method is its speed of computation.

The best combination of model parameters may be found using the R automatic procedure **`find.hyper()`** (Rigby and Stasinopoulos, 2005) which is based on the numerical optimisation function **`optim()`** (R Core Team, 2021) and selects hyperparameters as described in method 1. This was done for the BCPE distribution as well as the Normal, BCCG and BCT distributions.

Rigby and Stasinopoulos (2017) found that using a penalty of 3 was best when fitting a Box-Cox t distribution to head circumferences against age. The AIC was too "liberal" and the SBC too "conservative" in its choice of degrees of freedom but a penalty of 3 provided a good balance between the two. Stasinopoulos *et al*. (2017) also state that in practice the AIC leads to overfitting and the SBC underfitting and they recommend that a penalty between 2.5 and 4 works well. For a large sample size, a penalty of $\sqrt{\log(n)}$ may be used.

For a penalty of 3, the best fitting model was found to be the $BCPE(df_{\mu} = 6.3736, df_{\sigma} = 0\,, df_{\upsilon} = 3.3886, df_{\tau} = 0.6022, \xi = 0.0901)$ (model 3 in Table 4.3).

Table 4.3: GAMLSS Models with cubic spline smoothing for boys weight (penalty =3).

Distribution		**GAIC(3)**	$\boldsymbol{df_{\mu}}$	$\boldsymbol{df_{\sigma}}$	$\boldsymbol{df_{\nu}}$	$\boldsymbol{df_{\tau}}$	$\boldsymbol{\xi}$
1	NO	25402.34	8.6806	8.8234			0.2606
2	BCCG	25167.11	6.9234	0.7635	3.7965		0.2272
3	**BCPE**	**25080.94**	**6.3736**	**0**	**3.3886**	**0.6022**	**0.0901**
4	BCT	25083.45	7.3990	0	3.0950	0	0.1719

*****df of 0 means that the parameter is modelled linearly with age.

The diagnostics of the normalised quantile residuals of the model are displayed in Figure 4.15. There appears to be a random scatter in the plots of residuals vs fitted values and against age and the Normal Q-Q plot does not indicate any extreme outliers.

Figure 4.16 displays further model diagnostics in terms of a worm plot. The age range is split into 16 intervals and a detrended Normal Q-Q plot is plotted for each interval. This allows for identification of problems with model fit within each of the age intervals. The worm plots do not suggest any major inadequacies with the model fit as the worms are fairly horizontal and the majority of the points lie within the 95% confidence bands.

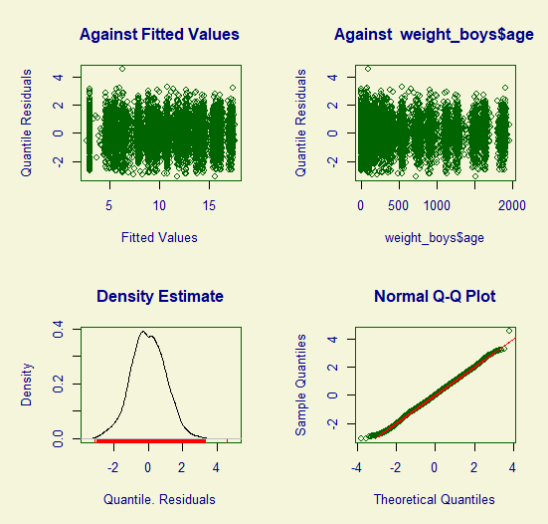


Figure 4.15: Residual diagnostics of BCPE model with cubic spline smoothing.

Analysis of the Q-statistics in Figure 4.17 revealed only one age interval for the σ parameter and one for the skewness parameter with |Z values| >2. The model fit was therefore deemed adequate.

The corresponding centiles are displayed in Figure 4.18. The proportion of observations falling below each centile for the fitted model and the sample (Table 4.4) are acceptably close and the curves are sufficiently smooth.

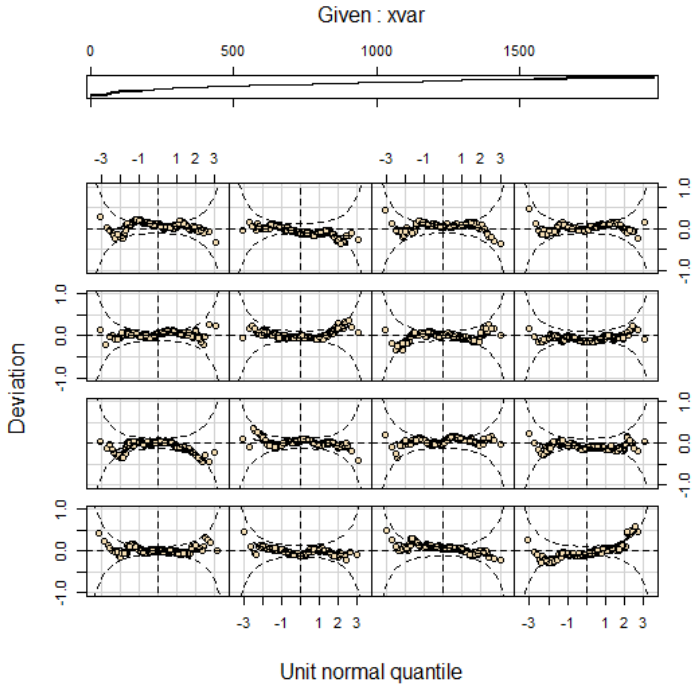


Figure 4.16: Worm plot of BCPE model with cubic spline smoothing.

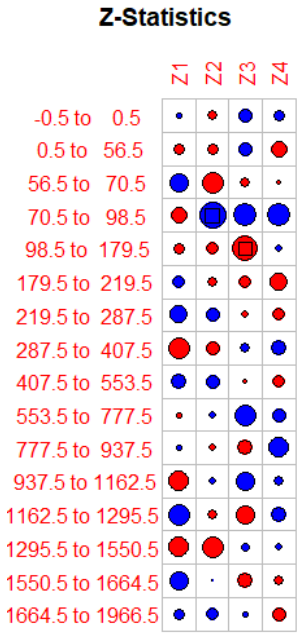


Figure 4.17: Q-statistics of BCPE model with cubic spline smoothing.

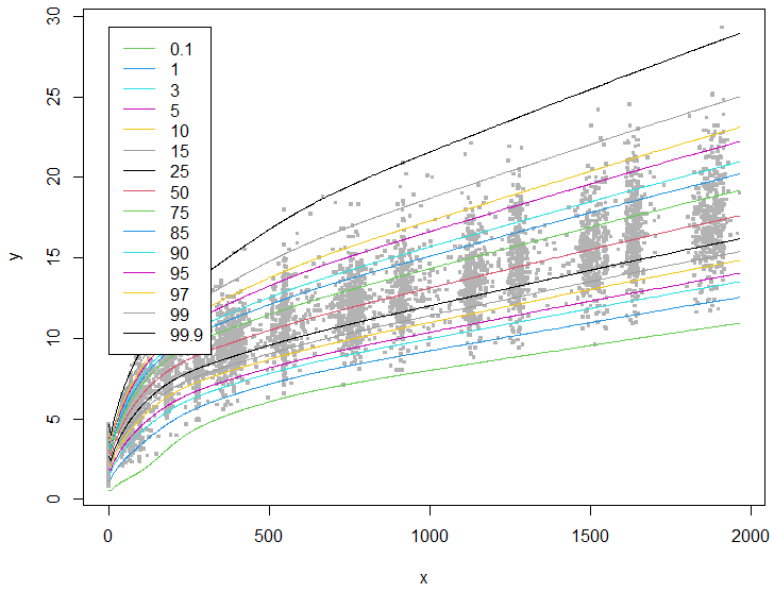


Figure 4.18: Centile curves for BCPE model with cubic spline smoothing.

Table 4.4: Proportion of observations below each centile curve for BCPE model with cubic spline smoothing.

Fitted centile	Proportion of observations falling below fitted centile
0.1	0.0278319
1	1.238519
3	3.256332
5	5.051489
10	9.212357
15	14.26385
25	24.50598
50	50.40356
75	74.68689
85	84.91511
90	89.84136
95	95.04592
97	97.23073
99	98.98414
99.9	99.90259

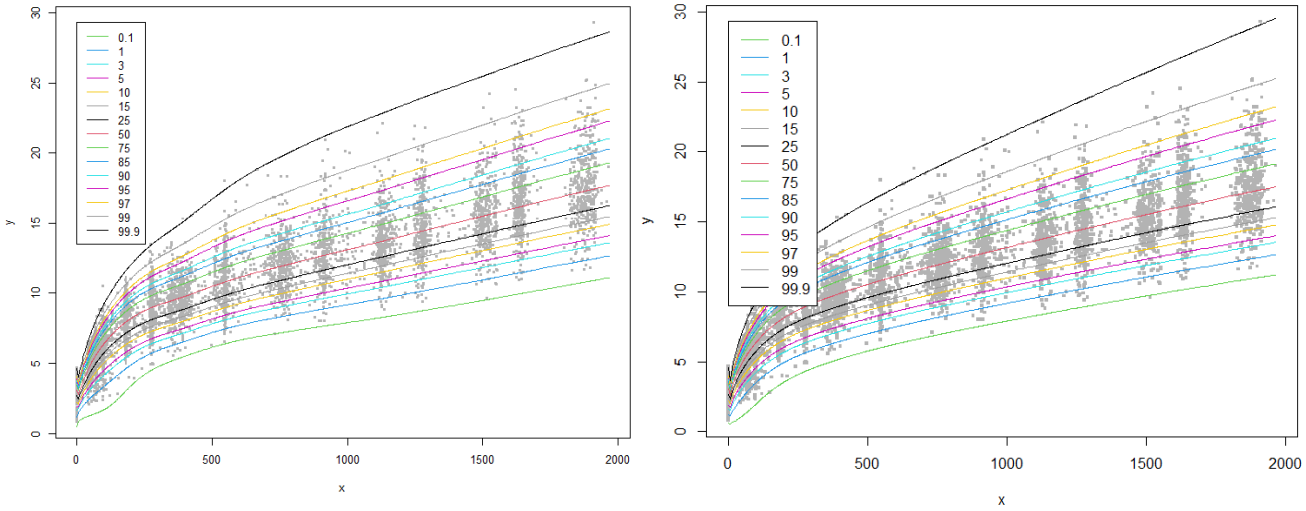
The sensitivity of the models and resulting centiles to different penalties was investigated. A higher penalty acts to reduce the degrees of freedom and thereby increase the degree of smoothing. Table 4.5 below lists the degrees of freedom for BCPE models with different penalties applied to show the sensitivity of the models to different penalties. Penalties of 2 (AIC), 3 and *log(n)* (equivalent to the Bayesian Information Criterion) were compared.

Table 4.5: Sensitivity of GAMLSS model (with cubic spline smoothing) parameters to different penalty values.

Penalty	df_{μ}	df_{σ}	df_{ν}	df_{τ}	**power**
2	9.4646	0.6118	3.9577	3.6329	0.1386
3	6.3736	0	3.3886	0.6022	0.0901
***log(n)* = 8.9174**	4.9367	0	1.1563	0	0.0563

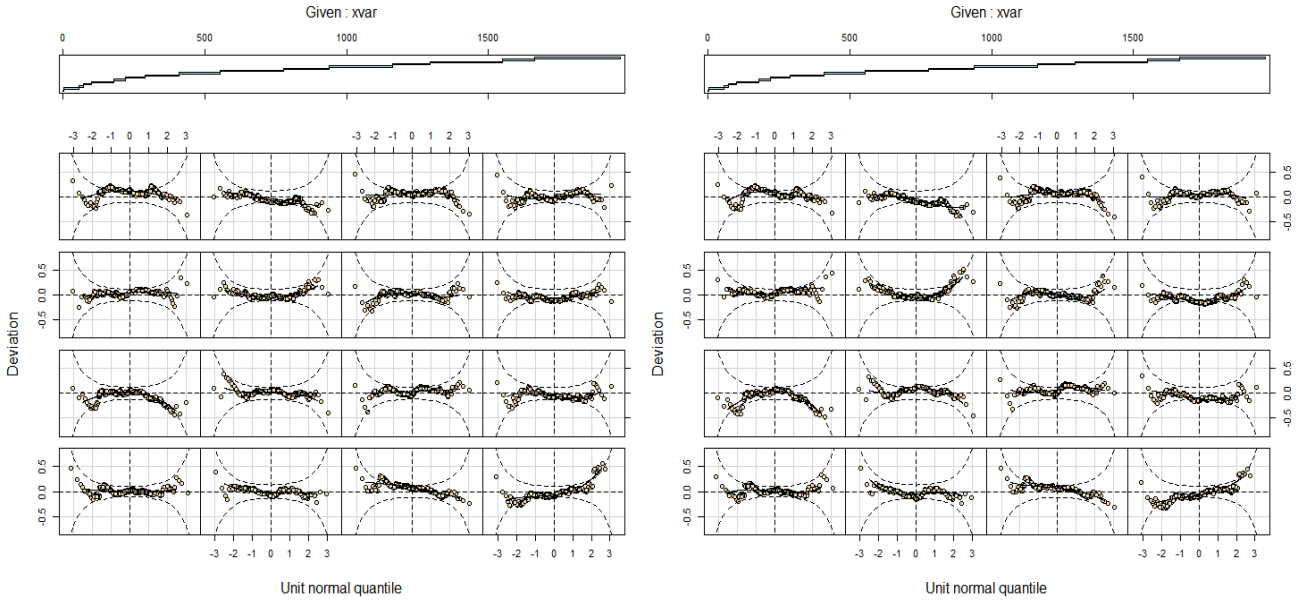
**df*=0 means relationship is linear with age

Figure 4.19 illustrates the trade-off between model fit and smoothness of the centile curves for different penalties. Using a penalty of 2 results in slightly wavy curves. The impact on model fit of increasing the penalty to *log(n)* (almost 9 in this case) is illustrated by the worm plots in Figure 4.20. Those for the more restrictive penalty deviate more from the desired horizontal form of the curves with more points lying outside of the 95% confidence bands. A penalty of 3 is a compromise between the two.



(a) **(b)**

Figure 4.19: Centile curves for penalties of (a) 2 and (b) *log(n)* for GAMLSS model with cubic spline smoothing.



(a) **(b)**

Figure 4.20: Worm plots for penalties of (a) 2 and (b) *log(n)* for GAMLSS model with cubic spline smoothing.

4.1.3.2 WHO procedure for selection of hyperparameters

The WHO chose to use the BCPE distribution to construct their growth curves which gave them the flexibility to account for skewness and kurtosis where necessary. For fixed values of 1 and 2 for ν (the skewness parameter) and τ (the kurtosis parameter) respectively, the BCPE distribution simplifies to the normal distribution. For $\nu \neq 1$ and τ equal to 2, the distribution is the same as the Box-Cox Cole and Green or the Box-Cox Normal distribution. Thus, by selecting the BCPE method, the WHO were able to compare three distributions: Normal, BCCG and BCPE. The WHO compared two smoothing methods – cubic smoothing splines and fractional polynomials – and cubic smoothing splines were chosen because they were found to be more flexible in all cases.

The WHO describe their model selection process as an add-up stepwise process starting from the simplest models and increasing in complexity.

1. The first step in the model selection process was to choose an appropriate age transformation to stretch the age scale for closer to birth because this is when the fastest growth occurs and compress it when growth is slow. This provides a uniform growth rate on the transformed age scale. Without this, models struggle to fit early infancy growth with reasonable degrees of freedom. When the degrees of freedom are increased too much, it results in too little smoothing at higher ages.
   Start with an arbitrary base model and compare models using a range of different transformation powers based on the global deviance.
2. Find the best values for df_μ and df_σ using the power found in step 1 and fixed values for ν and σ. A search was done for the best combination of df_μ and df_σ from predetermined ranges.
3. Search for the best value for df_ν, holding τ constant and using the values of df_μ and df_σ found in step 2. df_ν values in a predetermined range were tried and compared.
4. Check if the values of df_μ and df_σ can be improved upon by using the newly found df_ν, τ=2 and the power found in step 1.
5. Search for the best value of df_τ using values found in the above steps.
6. Check if the values of df_μ and df_σ can be improved upon by using the newly found df_τ and the power found in step 1.

7. Check for further improvement in transformation power given degrees of freedom values found in steps 2-5 above.

The starting model *BCPE(df_μ =10, df_σ =5, df_ν=5, df_τ=5)* was used as the base model for finding the best transformation power. Values between 0.1 and 1 in increments of 0.05 for the power parameter were tested and the chosen power of 0.15 had the smallest global deviance as can be seen in Table 4.6.

Table 4.6: Global deviances for different power parameters.

Power	GD	Power	GD
0.1	25002.61	0.6	25012.16
0.15	**25002.40**	0.65	25011.53
0.2	25002.82	0.7	25010.46
0.25	25003.96	0.75	25009.91
0.3	25005.52	0.8	25011.83
0.35	25007.18	0.85	25019.95
0.4	25008.76	0.9	25039.92
0.45	25010.21	0.95	25080.88
0.5	25011.37	1	25156.81
0.55	25012.08		

Next, a search for the best df_μ and df_σ was done, fixing the power parameter to 0.15. As done by the WHO, ν and τ were fixed at 1 and 2 respectively. The best combination of df_μ and df_σ was selected from the ranges df_μ = *{4, 5, 6, , …, 15}* and df_σ = *{0, 1, 2, …, 10}* based on the GAIC measure with a penalty of 3. df_μ =9 and df_σ =10 was the best combination and the resulting model was Model 1 presented below that will now be further improved by choosing the optimal skewness parameters. Based on the AIC, the chosen model would have been df_μ =15 and df_σ =10 and based on the BIC, it would have been df_μ =5 and df_σ =0. Using GAIC(6), the best combination would be df_μ =6 and df_σ =0. An extract of the GAIC(#) values resulting from the different combinations of df_μ and df_σ is shown in Table 4.7.

Model 1: ***BCPE(x=age^0.15, df_μ =9, df_σ=10, ν =1, τ =2).***

Table 4.7: GAIC(#) comparison for different combinations of df_μ and df_σ.

df_μ	df_σ	AIC	GAIC(3)	GAIC(6)	BIC	Total df
6	**8**	25385.98	25403.98	25457.99	25509.83	14
	9	25384.73	25403.73	25460.74	25515.47	15
	10	25383.37	25403.38	25463.39	25520.99	16
7	**8**	25383.88	25402.88	25459.88	25514.60	15
	9	25382.64	25402.64	25462.64	25520.24	16
	10	25381.27	25402.27	25465.27	25525.75	17
8	**8**	25382.53	25402.53	25462.54	25520.13	16
	9	25381.28	25402.28	25465.29	25525.76	17
	10	25379.91	25401.91	25467.91	25531.27	18
9	**8**	25381.55	25402.55	25465.56	25526.04	17
	9	25380.29	25402.29	25468.30	25531.67	18
	10	25378.90	**25401.90**	25470.91	25537.16	19

The residual diagnostics for Model 1 are shown in Figure 4.21. There is a random scatter in the plot of residuals vs fitted values. The Normal Q-Q plot, however, shows departures from normality at the tails. The worm plots in Figure 4.22 indicate the following:

- Reasonable fit is obtained for the μ and σ curves as the worms do not pass significantly above or below the origin, nor do they have a significantly positive or negative slope.
- There is clear evidence that skewness is not adequately accounted for because the worms are either U-shaped or inverted U-shaped for most of the age ranges.
- A few age ranges have worms with an S curve shape, suggesting residual kurtosis.

The above observations are confirmed by the Q-statistics (Figure 4.23) which show that the absolute values of Z3 (skewness) and Z4 (kurtosis) mostly exceed 2.

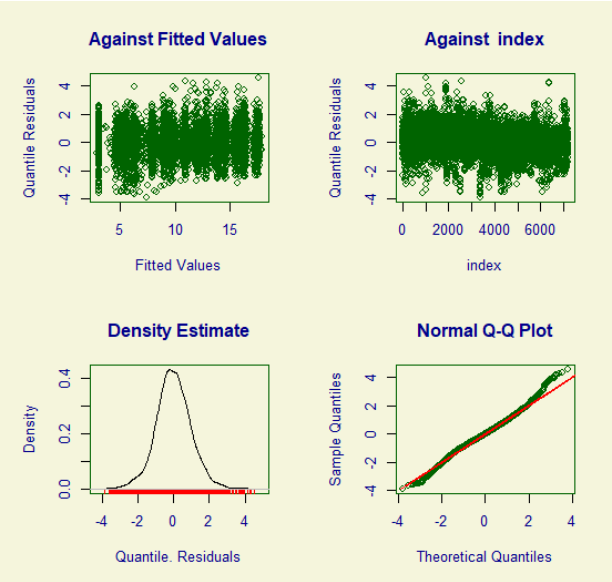


Figure 4.21: Residual diagnostics for Model 1: *BCPE(x=age^0.15, df_μ =9, df_σ=10, ν =1, τ =2).*

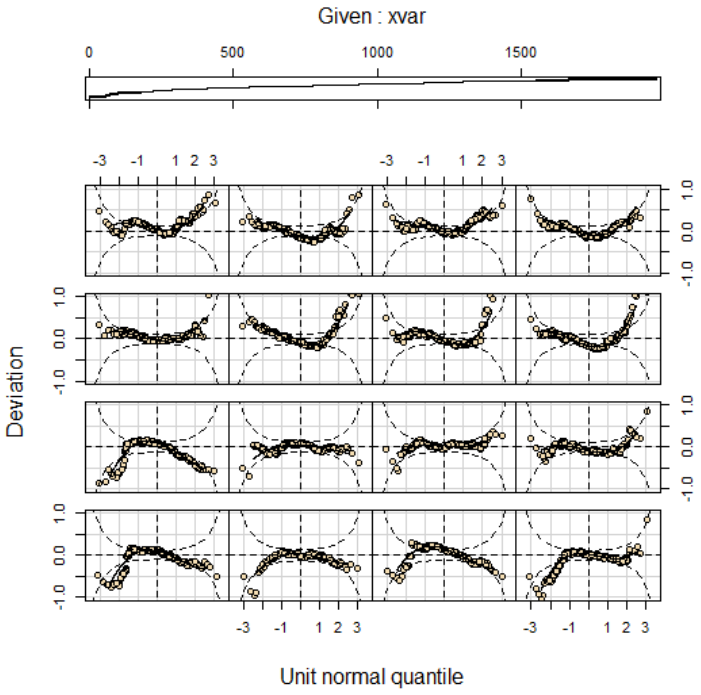


Figure 4.22: Worm plot for Model 1: *BCPE(x=age^0.15, df_μ =9, df_σ=10, ν =1, τ =2).*

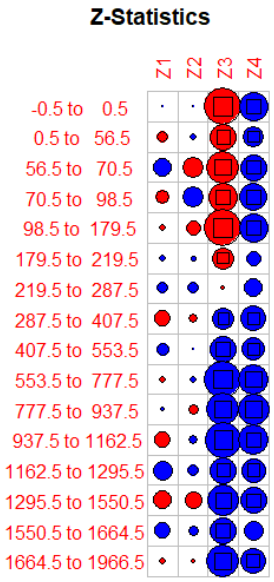


Figure 4.23: Q-statistics for Model 1: *BCPE(x=age^0.15, df_μ =9, df_σ=10, ν =1, τ =2).*

The residual skewness prompted a search for the skewness degrees of freedom parameter which results in the smallest GAIC(3) for the model *BCPE(x=age^0.15, df_μ =9, df_σ=10, df_ν=?, τ =2).* The GAIC(3) values for a range of df_ν = *{0, 1, 2, …, 7}* are shown in Table 4.8 and the value of 4 was found to be best, which gave rise to Model 2.

Model 2: *BCPE(x=age^0.15, df_μ =9, df_σ=10, df_ν=4, τ =2).*

Table 4.8: GAIC(#) comparison for different values of df_ν.

df_ν	AIC	GAIC(3)	GAIC(6)	BIC	Total df
0	25162.92	25189.09	25267.62	25342.99	19
1	25162.92	25189.09	25267.62	25342.99	20
2	25154.49	25181.49	25262.50	**25340.27**	21
3	25149.98	25177.98	**25261.99**	25342.64	22
4	**25148.79**	**25177.80**	25264.81	25348.33	23
5	25149.22	25179.22	25269.24	25355.65	24
6	25150.17	25181.17	25274.19	25363.48	25
7	25151.18	25183.18	25279.19	25371.36	26

Model 2 has an improved fit compared to Model 1. There is no longer a significant deviation from normality in the upper tail on the Normal Q-Q plot in Figure 4.24. The worm plots in Figure 4.25 indicate a better fit compared to Model 1 even though a few points still lie outside of the 95% confidence bands. There are no longer any worms which are very obviously U-shaped, however a number of the worms have a slight S-shape which is indicative of residual kurtosis in the data.

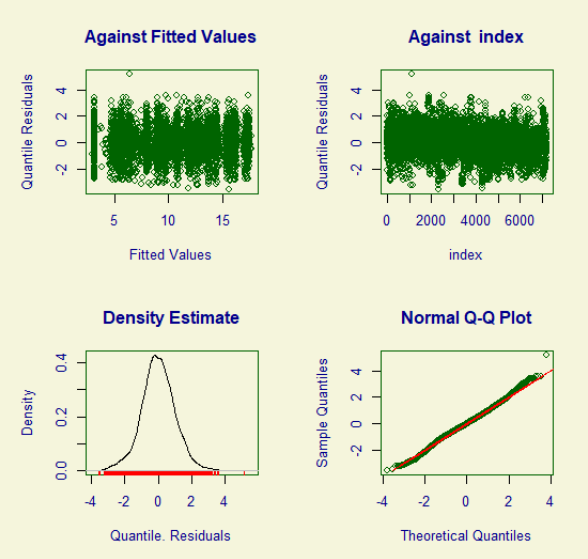


Figure 4.24: Residual diagnostics for Model 2: *BCPE(x=age^0.15, df_μ =9, df_σ=10, df_ν=4, τ =2).*

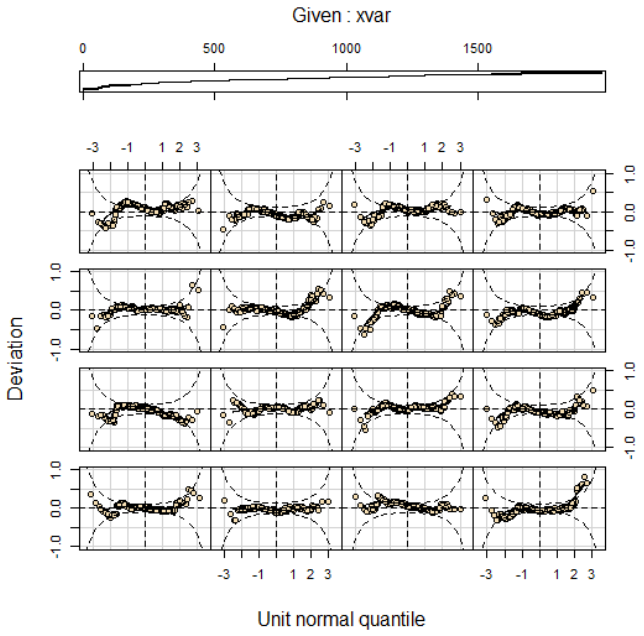


Figure 4.25: Worm plot for Model 2: *BCPE(x=age^0.15, df_μ =9, df_σ=10, df_ν=4, τ =2).*

The Q-statistics (Figure 4.26) confirm the improved fit and imply the need to model kurtosis. The absolute values of the Q-statistics for the mean, variance and skewness parameters are now almost all less than 2, while z4 has a number of age ranges for which the values exceed 2.

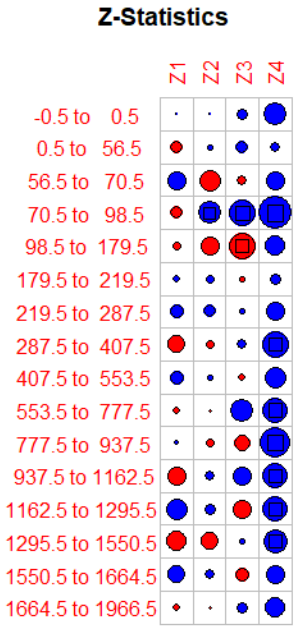


Figure 4.26: Q-statistics for Model 2: *BCPE(x=age^0.15, df_μ =9, df_σ=10, df_ν=4, τ =2).*

Before modelling kurtosis, another search was done for df_μ and df_σ to see if the model could be improved further after modelling skewness as a cubic spline with age. A GAIC(#) comparison for an extract of the different combinations is shown in Table 4.9. Based on the GAIC(3), df_μ =7 and df_σ =1 was the preferred combination, producing Model 3 below.

Model 3: *BCPE(x=age^0.15, df_μ=7, df_σ=1, df_ν=4, τ =2).*

Table 4.9: GAIC(#) comparison for different combinations of df_μ and df_σ after setting df_ν.

df_μ	df_σ	AIC	GAIC(3)	GAIC(6)	BIC
5	**1**	25156.99	25173.14	25221.60	25268.12
	2	25158.42	25175.42	25226.42	25275.37
	3	25159.48	25177.48	25231.47	25283.30
	4	25159.75	25178.75	25235.75	25290.47
6	**1**	25151.56	25168.72	25220.18	25269.58
	2	25152.99	25170.99	25224.99	25276.83
	3	25154.02	25173.02	25230.02	25284.74
	4	25154.30	25174.30	25234.30	25291.90
7	**1**	25149.57	**25167.72**	25222.18	25274.47
	2	25150.98	25169.98	25226.99	25281.71
	3	25152.00	25172.00	25232.00	25289.60
	4	25152.27	25173.27	25236.27	25296.74
8	**1**	25148.69	25167.84	25225.31	25280.47
	2	25150.10	25170.10	25230.10	25287.71
	3	25151.11	25172.11	25235.11	25295.59
	4	25151.35	25173.35	25239.35	25302.71

The new values of df_μ and df_σ did not have a material impact on the model fit as seen in the worm plot and Q-statistics. Models with different values for df_τ in the range of 1 to 7 were then compared and the GAIC(#) values are shown in Table 4.10. The GAIC(3) criterion favoured df_τ =1, resulting in Model 4 below.

Model 4: ***BCPE(x=age^0.15, df_μ=7, df_σ=1, df_ν=4, df_τ =1).***

Table 4.10: Comparison of models based on different df_τ values.

df_τ	AIC	GAIC(3)	GAIC(6)	BIC/SBC	Total df
1	25065.50	**25086.81**	**25150.73**	**25212.09**	13
2	25065.94	25088.10	25154.56	25218.37	14
3	25065.30	25088.45	25157.92	25224.61	15
4	**25064.94**	25089.10	25161.58	25231.15	16
5	25065.51	25090.67	25166.16	25238.62	17
6	25066.59	25092.76	25171.26	25246.61	18
7	25067.84	25095.01	25176.50	25254.73	19

Finally, using the degrees of freedom parameters found in the above steps, a new search was done to see if a different age transformation could improve the model fit. The power resulting in the lowest global deviance was 0.10 which is only marginally lower than the initial power of 0.15 and the impact on model fit and the resulting centiles was negligible so Model 4, *BCPE(x=age^0.15,* df_μ*=7,* df_σ*=1,* df_ν*=4,* df_τ *=1)* was selected as the final model. Since the degrees of freedom for the variance and kurtosis parameters are so low, the model could be simplified by setting these parameters to change linearly with age. The resulting model would be very similar to that found using the automatic procedure. The worm plot and Q-statistics in Figures 4.27 and 4.28 respectively both show that the model fits the data well.

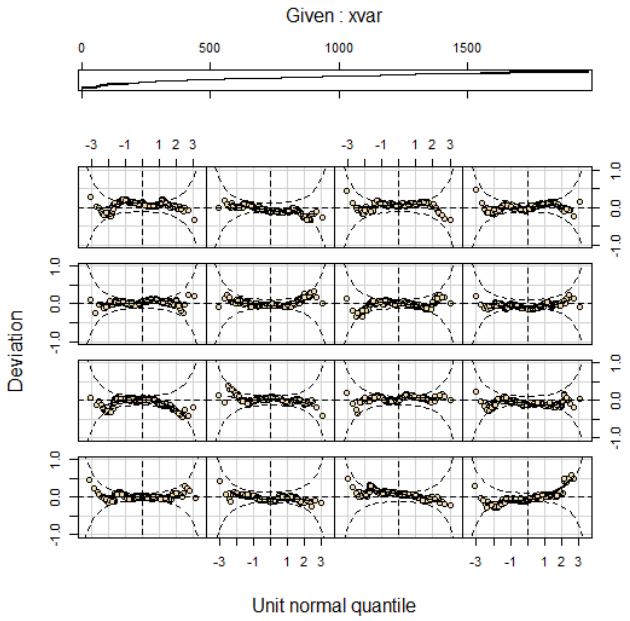


Figure 4.27: Worm plot for Model 4: *BCPE(x=age^0.15,* df_μ*=7,* df_σ*=1,* df_ν*=4,* df_τ *=1).*

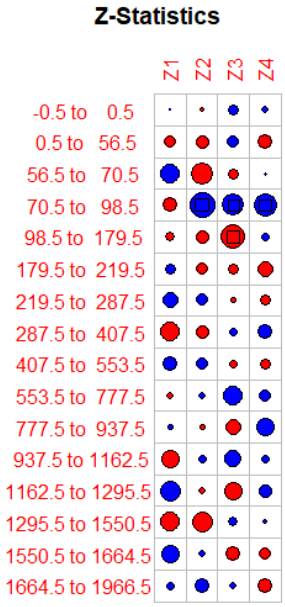


Figure 4.28: Q-statistics for Model 4: *BCPE(x=age^0.15, df_μ=7, df_σ=1, df_ν=4, df_τ =1).*

The fitted models for the parameters μ, σ, ν and τ are displayed in Figure 4.29. The plot for μ shows that the rate of weight gain for children is highest in the first year of life and tapers off with increasing age, becoming almost linear from age 2 to 5 years. The σ plot indicates that variation in weight is high at birth and falls quickly until approximately age 1, after which it decreases more gradually until age 5. The skewness parameter starts high at birth indicating negative skewness, decreases steadily until age 2.5 and thereafter decreases very slowly until age 5. The kurtosis parameter also starts high and decreases gradually from birth to age 5.

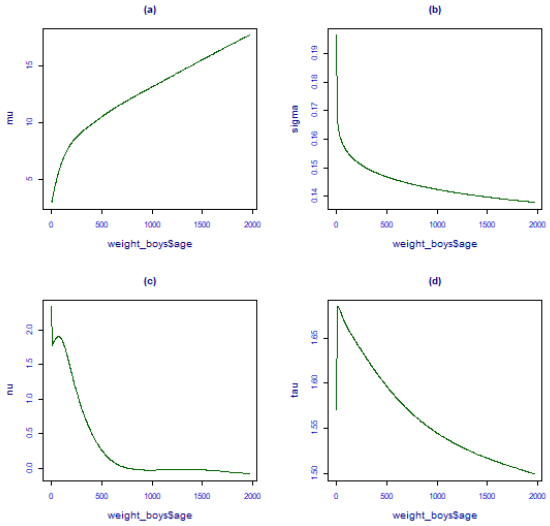


Figure 4.29: Fitted parameter curves for Model 4: *BCPE(x=age^0.15, df_μ=7, df_σ=1, df_ν=4, df_τ =1).*

The centile curves for Model 4, plotted in Figure 4.30, are smooth and follow the expected pattern based on the data. Table 4.11 also show that there is also close agreement between the proportions of observed weight measurements below each centile curve and the expected proportions based on the fitted models.

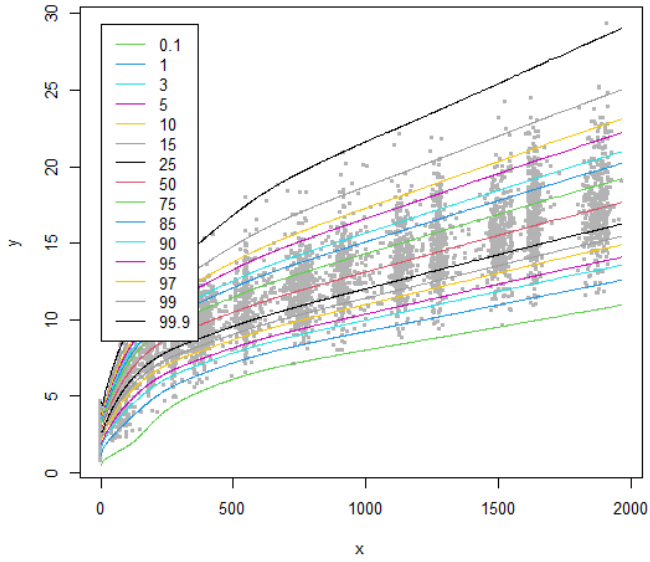


Figure 4.30: Centile curves for Model 4: *BCPE(x=age^0.15, df_μ=7, df_σ=1, df_ν=4, df_τ =1).*

Table 4.11: Proportion of observations below each centile curve for Model 4.

Fitted centile	Proportion of observations falling below fitted centile
0.1	0.01391595
1	1.196772
3	3.284164
5	5.065405
10	9.128862
15	14.26385
25	24.53382
50	50.32007
75	74.63123
85	84.88728
90	89.84136
95	95.10159

97	97.27247
99	98.99805
99.9	99.90259

4.1.4 Final parametric/semiparametric model

Table 4.12 compares the fitted parametric and semi-parametric models based on the GAIC(#) using various penalties. The model $BCPE(df_\mu = 6.3736, df_\sigma = 0, df_\upsilon = 3.3886, df_\tau = 0.6022, \xi = 0.0901)$ found using the automatic procedure **`find.hyper()`** (Rigby and Stasinopoulos, 2005) in RR obtained the lowest GAIC for penalties 2 (AIC), 3, 6 and *log(n)* (BIC). However, since the resulting model of the WHO model selection process was very similar to this, it was the preferred final model. This was the model *BCPE(x=age^0.15,* df_μ*=7,* df_σ*=1,* df_ν*=4,* df_τ *=1).*

Table 4.12: Comparison of best models found for each parametric/semi-parametric modelling approach for boys' weight.

Model	AIC	GAIC(3)	GAIC(6)	BIC	df
BCPE with fractional polynomials "smoothing"	25094.64	25122.64	25206.64	25287.28	28.0000
BCPE with Berkey-Reed Second order model for μ and cubic smoothing splines for σ, ν and τ	25146.09	25169.10	25238.11	25304.35	23.0026
BCPE with cubic spline smoothing – auto procedure	**25062.57**	**25080.94**	**25136.04**	**25188.93**	**18.3665**
BCPE with cubic spline smoothing – WHO process	**25065.50**	**25086.81**	**25150.73**	**25212.09**	**21.3064**

4.1.5 Quantile Regression

Model construction involved selecting a set of knots to be used to construct the basis splines. Models may also be specified in terms of the degrees of freedom. The number of knots is related to the degrees of freedom: the degrees of freedom are equal to $d+K$ where d is the degree of the B-splines and K is the number of interior knots. Specifying a degree of 3 means that a cubic B-spline basis will be used. The recommendations of Wei *et al.* (2006) to use fairly evenly spaced knots with extra knots during periods of rapid growth such as in early infancy, and to keep the set of knots the same for different centiles were followed.

The plots in Figure 4.31 show the B-spline basis functions constructed for the following sets of knots in addition to the boundary knots of 0 and 1966 days.

(a) 14 weeks, 48 months (degrees of freedom = 5).
(b) 10 weeks, 6 months, 18 months, 48 months (degrees of freedom = 7).
(c) 6 weeks, 14 weeks, 6 months, 18 months, 36 months, 48 months (degrees of freedom = 9).
(d) 6 weeks, 10 weeks, 14 weeks, 6 months, 12 months, 18 months, 24 months, 30 months, 36 months, 42 months, 48 months, 54 months, 60 months (degrees of freedom = 16).

The corresponding growth reference curves are presented in Figure 4.32.

The higher the degrees of freedom, the more flexible the quantile regression function can be. Plot (d) in Figure 4.31 shows the basis functions for knots placed at each scheduled visit and, from the corresponding plot of the growth reference curves in Figure 4.32, it is obvious that the model is overfitting the data. Van Buuren (2007) explains that evidence of this can be seen in the worm plot where less than 1% of the points fall outside the 95% confidence bands where we expect 5% under the true model.

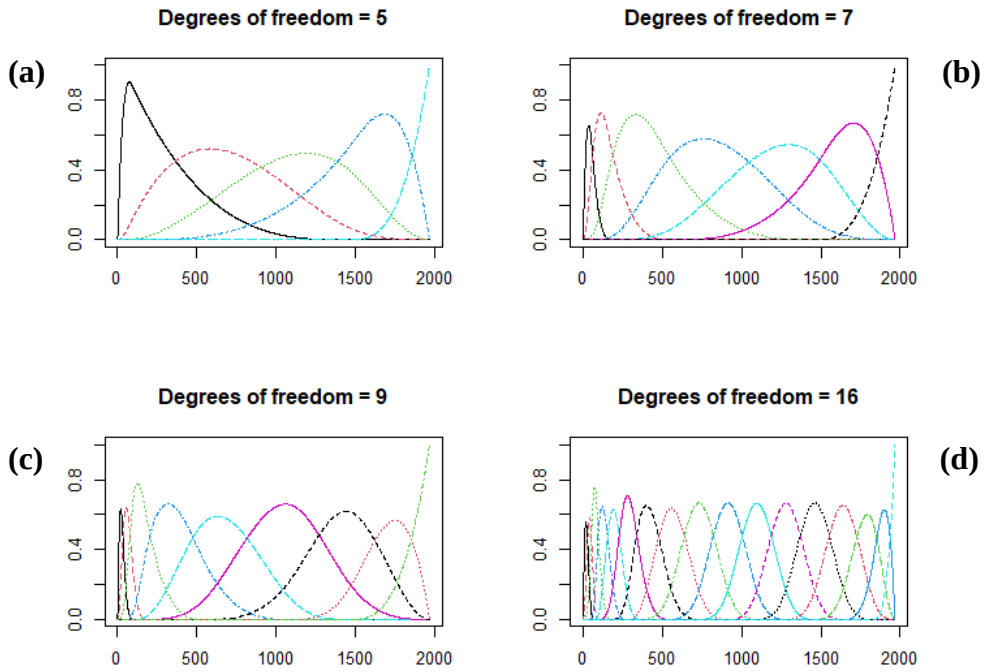


Figure 4.31: B-spline basis function for Drakenstein boys.

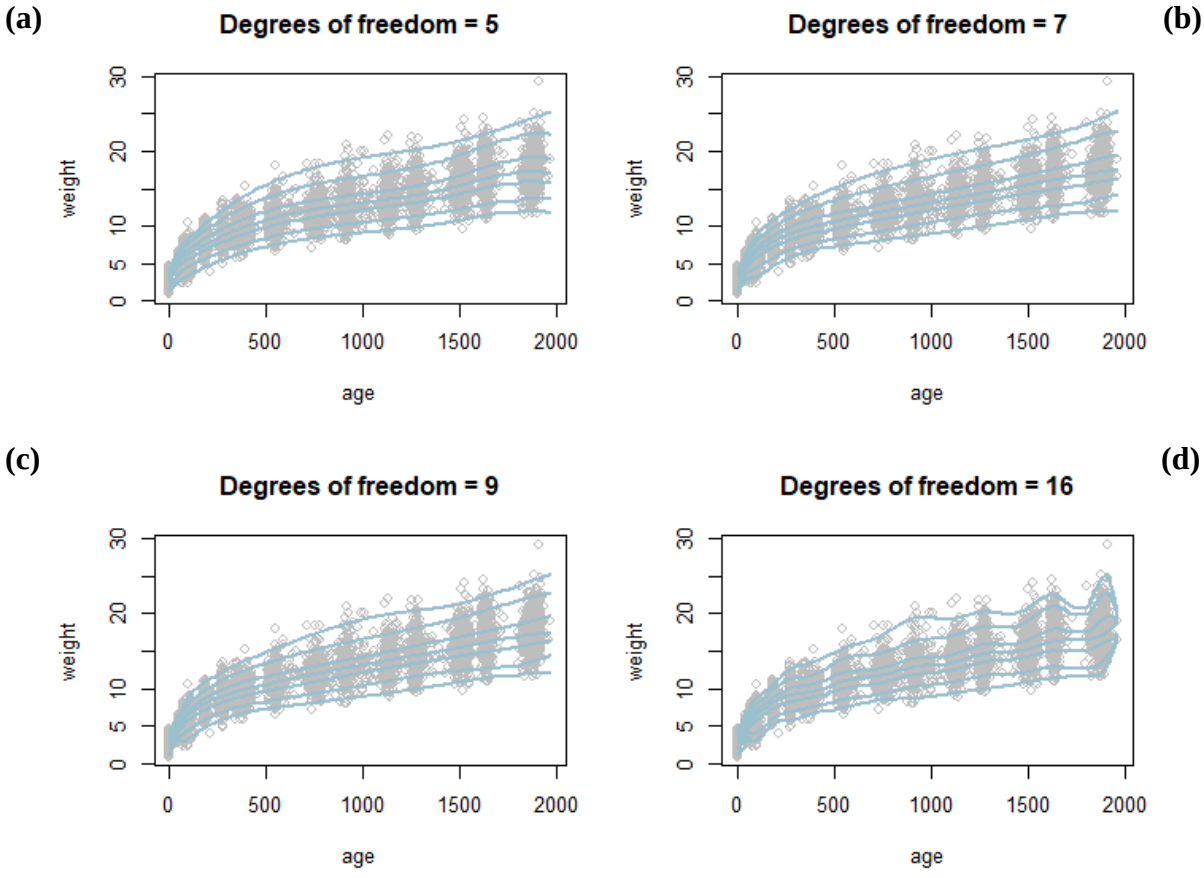


Figure 4.32: Centile curves corresponding to various B-spline basis functions.

Worm plots were constructed for the quantile regression models as described in van Buuren (2007) and used to assess model fit and choose the best model. Quantile regression models were fitted for various values of τ. The same grid of τ values as that used by van Buuren (2007) was used: $\tau \in \{0.001, 0.002, \dots, 0.01, 0.02, \dots, 0.98, 0.99, 0.991, 0.992, \dots, 0.999\}$. This resulted in k=117 models being fitted. The fitted quantile values were used to calculate approximate probability density functions (PDF) and cumulative distribution functions (CDF) values for the unknown distribution using the flexDist() function in R. The probability integral transform was then used to calculate the normalised quantile residuals which were used to produce the worm plot. Figures 4.33 to 4.36 show the worm plots constructed for quantile regression models based on 5, 7, 9 and 16 degrees of freedom respectively.

Degrees of freedom = 5

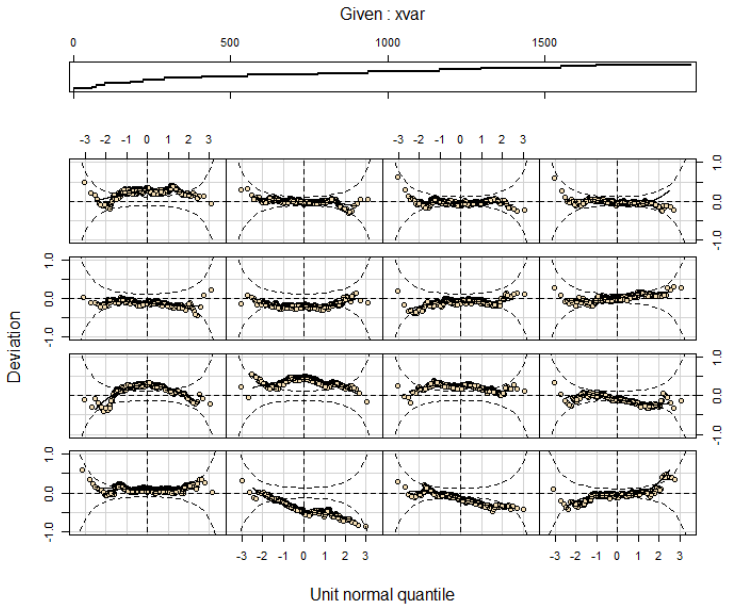


Figure 4.33: Centile curves for quantile regression model with 5 degrees of freedom.

Degrees of freedom = 7

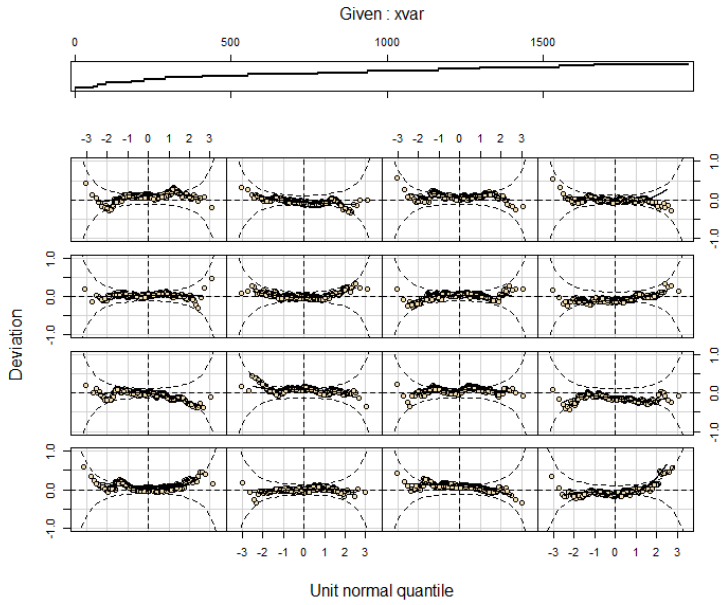


Figure 4.34: Centile curves for quantile regression model with 7 degrees of freedom.

Degrees of freedom = 9

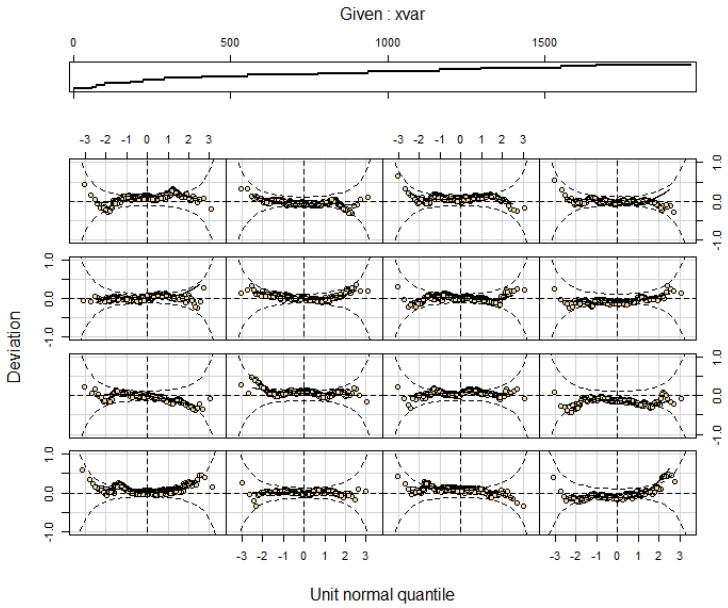


Figure 4.35: Centile curves for quantile regression model with 9 degrees of freedom.

Degrees of freedom = 16

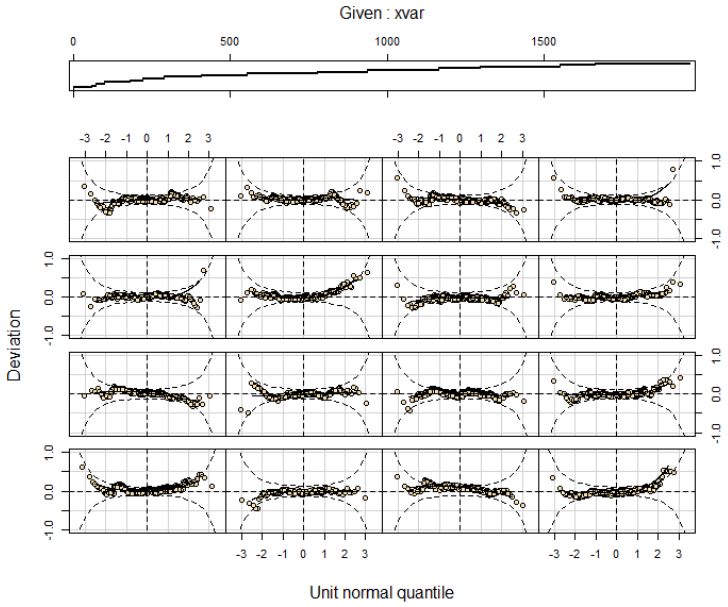


Figure 4.36: Centile curves for quantile regression model with 16 degrees of freedom.

The model with 16 degrees of freedom fits the data well as indicated by the flat worms. On the opposite extreme, the model with only 5 degrees of freedom shows clear signs of poor model fit: the curves for a number of age ranges pass though significantly above or below the origin and deviate from the desired flat shape.

Model b with 7 degrees of freedom was the chosen model. It adequately captured the patterns in the data without overfitting, producing nice smooth growth reference curves, and further increases in the degrees of freedom did not result in meaningful improvements in the worm plot.

The chosen quantile regression model was compared to the final selection of parametric/semi-parametric model which was the model *BCPE(x=age^0.15,* df_μ*=7,* df_σ*=1,* df_ν*=4,* df_τ *=1).* In Figure 4.37, the growth reference curves for the two models are plotted together. Both models provide a good fit to the data and there is close agreement between the curves, affirming the final semi-parametric model choice. The quantile regression curves do not highlight any locations where the data might deviate from the assumed model. Contrast this to Figure 4.38 where the quantile regression curves are overlaid on centile curves

constructed assuming a normal distribution – there are some noticeable deviations, particularly for the more extreme centiles.

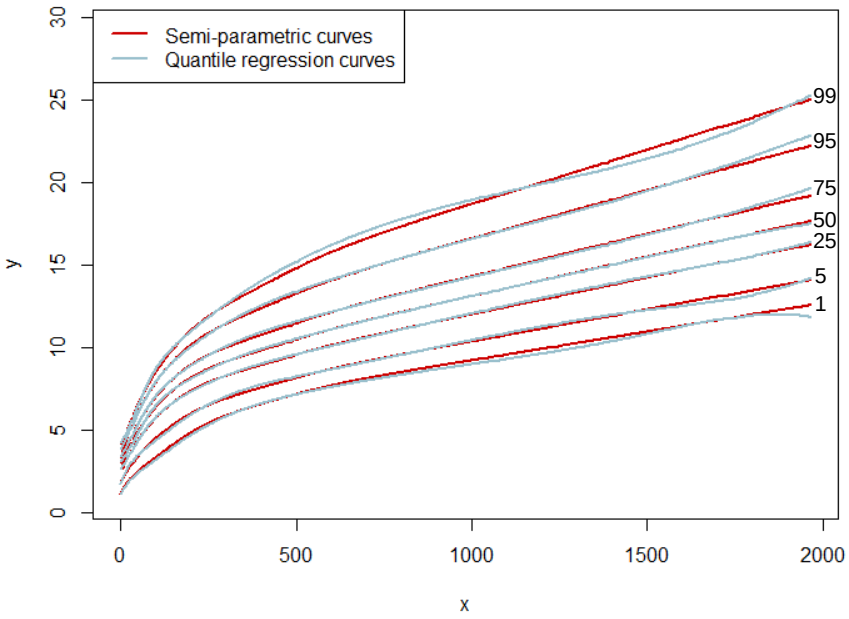


Figure 4.37: Comparison of final semi-parametric model for boys' weight with quantile regression model.

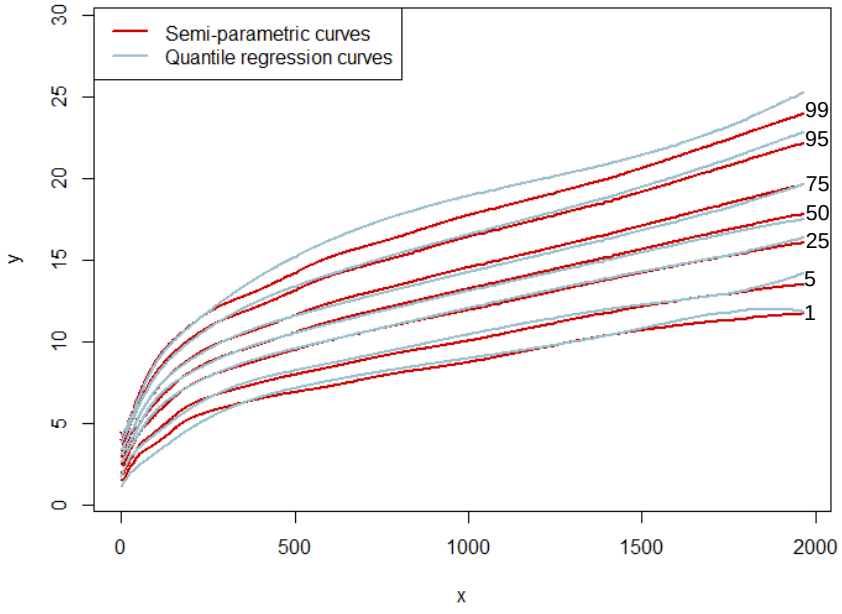


Figure 4.38: Comparison of centile curves derived using quantile regression and a model assuming a Normal distribution.

4.1.6 Confidence intervals for centiles

Confidence intervals may be calculated and plotted to provide a measure of uncertainly for the centile curves. It is possible to obtain the exact standard error curve for the 50th centile since the GAMLSS model provides estimates of the parameter curves with their corresponding standard error curves and the μ curve estimates the 50th centile. It is, however, not possible to obtain the exact standard error curves for the other centiles because the centile curves are non-linear functions of the moment curves (Cole, 2021).

In order to obtain the standard error curves, the advice of Cole (2021) was followed, who showed that the standard error curves may be obtained using bootstrapping. Samples of the data were repeatedly drawn with replacement, each time fitting the model to the new data set and saving the fitted parameter curves. These parameter estimates were then used to calculate the estimates of the centiles at each age and the standard deviation of these centile estimates is an estimate of the standard error curve at a particular age. These estimates together form the standard error curve for the centile which are then used to determine the confidence intervals for the curves.

The 1st, 25th, 50th, 75th and 99th centile curves together with their confidence intervals for the final semi-parametric model for boys' weight are plotted in Figure 4.39. The intervals are wider for the more extreme centiles, particularly for the higher centiles. Cole (2021) states that the standard errors for centiles above the median are larger than those below simply because the centile itself is larger. Overall, the confidence bands for the centile curves are fairly narrow, implying that the derived centile curves provide a precise representation of the growth centiles for the Drakenstein boys.

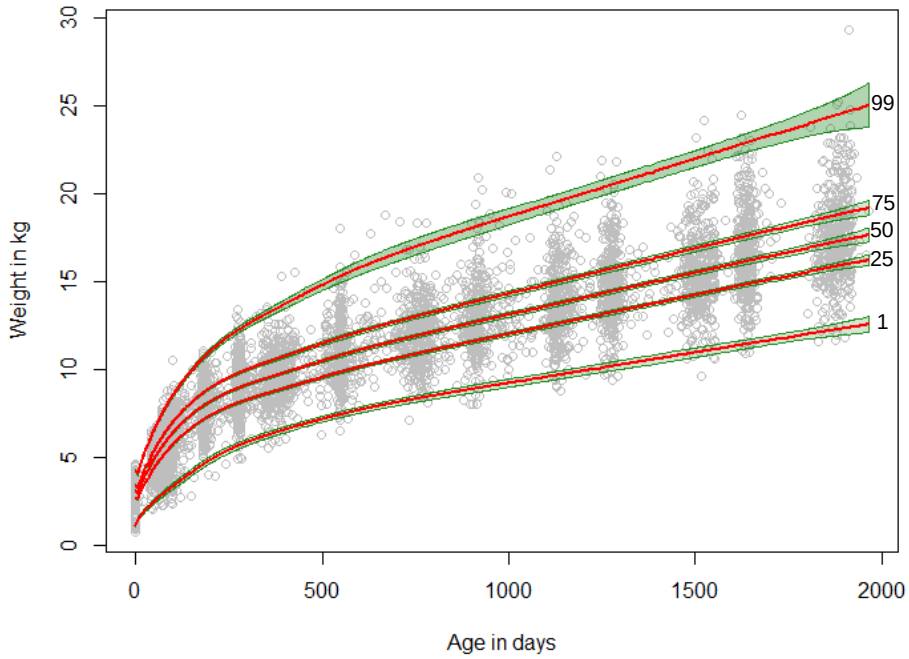


Figure 4.39: Centile curves and their confidence intervals for final semi-parametric model for boys' weight.

4.2 Weight-for-age: girls

In the interest of consistency, the final chosen modelling approach applied to boys' weights was also used to derive the centile curves for girls. This was to use the GAMLSS framework with the BCPE distribution and curve smoothing by cubic splines, with the smoothing parameters and transformation power chosen according to the WHO methodology. Density histograms of girls' weights plotted along with theoretical densities of the BCPE distribution at each visit suggested that the BCPE distribution will provide a good fit to the data.

The final model chosen was *BCPE(x=age^0.1,* df_μ*=6,* df_σ*=6,* df_ν*=3,* df_τ *=0)* with τ modelled linearly with age. The choice of df_ν=3 was based on the GAIC(6) because the GAIC(3) selected df_ν=9 which resulted in wavy centile curves and little difference in terms of model fit. The worm plot for the final model (Figure 4.40) shows reasonably good fit to the data as almost all worms are approximately horizontal. The Q-statistics in Figure 4.41 confirm this: for each of μ, σ, ν and τ there is only one age interval for which the absolute value of the Q-statistic exceeds 2.

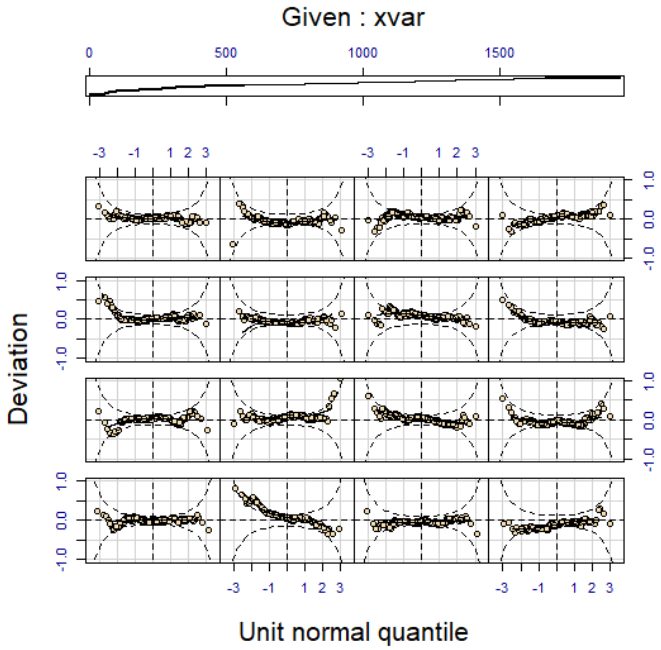


Figure 4.40: Worm plot of BCPE model with cubic spline smoothing for girls' weights.

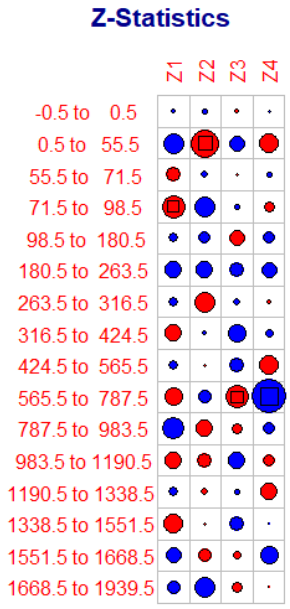


Figure 4.41: Q-Statistics for BCPE model with cubic spline smoothing for girls' weights.

The corresponding centile curves are shown in Figure 4.42. The curves are smooth and appear to follow the pattern of the observed data well and the observed and fitted proportions of cases below each centile (Table 4.13) are quite close.

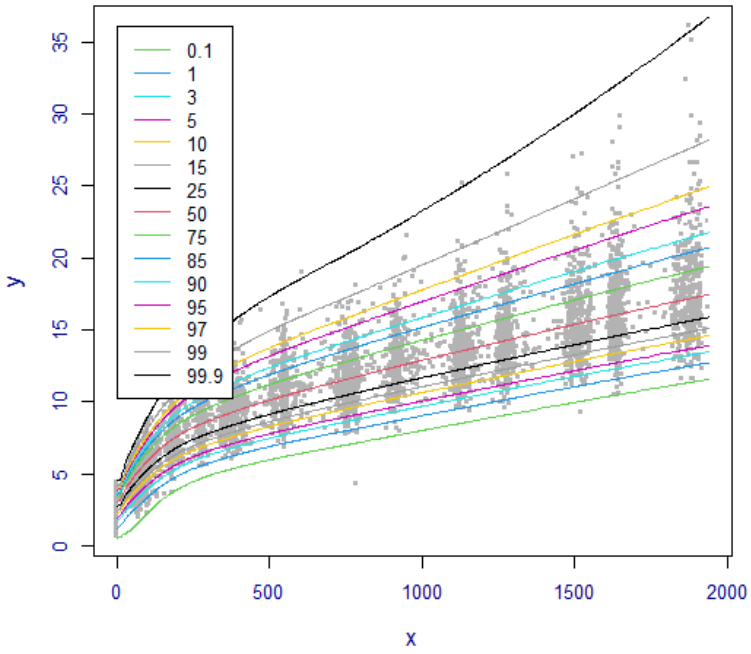


Figure 4.42: Centile curves for BCPE model with cubic spline smoothing for girls' weights.

Table 4.13: Proportion of observations below each centile curve

Fitted centile	Proportion of observations falling below fitted centile
0.1	0.03021148
1	0.9818731
3	2.809668
5	4.712991
10	9.924471
15	15.45317
25	25.54381
50	49.89426
75	74.92447
85	85.15106
90	90.10574
95	95.09063
97	97.02417
99	98.83686
99.9	99.90937

The same set of knots was used to fit the quantile regression model for boys' and girls' weights - 10 weeks, 6 months, 18 months, 48 months – which translates to a model with 7 degrees of freedom. The worm plot in Figure 4.43 indicates that the quantile regression model provides a good fit to the data.

The quantile regression centiles are overlaid on the centiles derived above using a semi-parametric model (Figure 4.44). There is very close agreement between the curves except for the upper ends of the 95^{th} and 99^{th} centiles where the quantile regression curves turn upwards slightly. It was previously discussed that quantile sheets regression was developed to solve the problems of crossing curves and irregular curves at the extremes of the data so this might result in an improvement. Despite the slight deviations, the curves were deemed adequately close to conclude that the chosen semi-parametric model provides a good fit to the data.

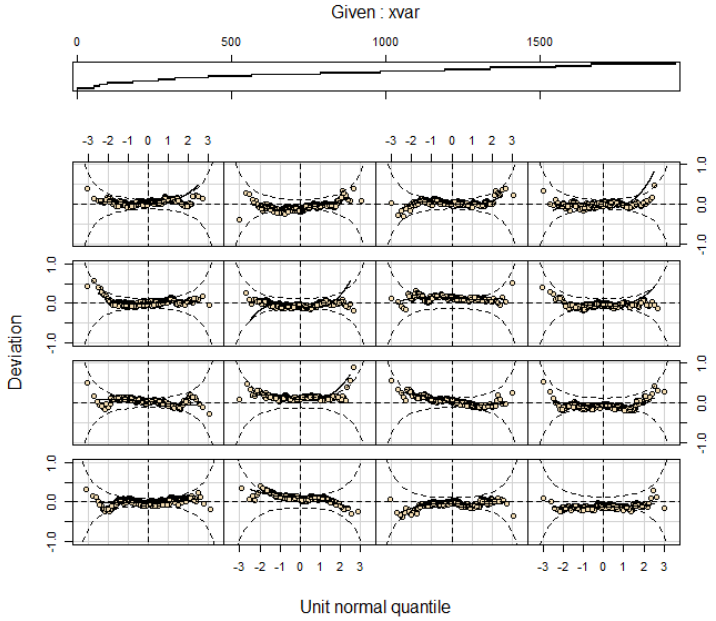


Figure 4.43: Worm plot for quantile regression model of girls' weights.

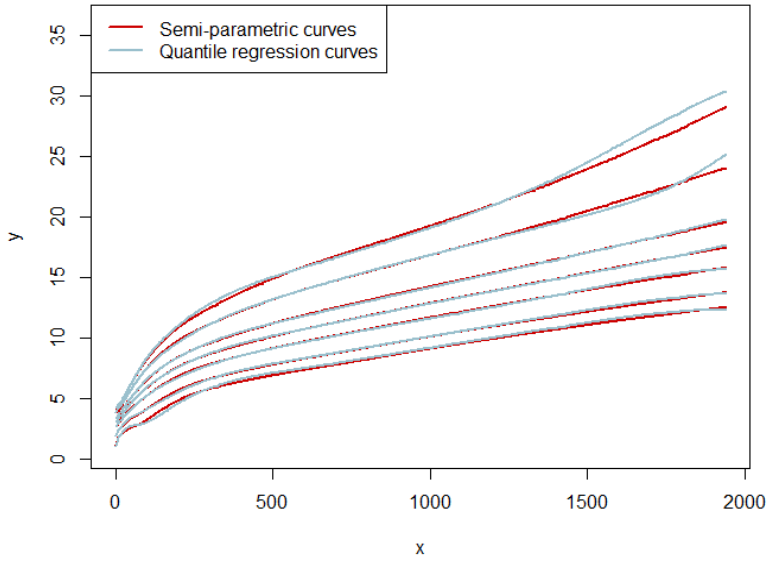


Figure 4.44: Quantile regression vs BCPE model with cubic spline smoothing for girls' weights.

4.3 Length/Height-for-age: boys

Based on plots of observed and theoretical distributions, the BCPE and BCT distributions appeared to be the most promising candidates for selection. For boys' heights, the BCT distribution was found to provide the best fit to the data when using both fractional polynomial and cubic splines smoothing. Since the height data exhibits less skewness than the weight data, the parametric approaches, fractional polynomials and Berkey-Reed models, provided a better fit to the data than was achieved for the weight data.

The final fractional polynomial model was

$$BCT(\mu = \beta_0 + \beta_1 \log(age) + \beta_2 \log(age) * \log(age) + \beta_3 age^2,$$
$$\sigma = \beta_0 + \beta_1 age + \beta_2 age * \log(age) + \beta_3 age * \log(age)^2,$$
$$\nu = \beta_0 + \beta_1 age^{-0.5} + \beta_2 age^{-0.5} * \log(age) + \beta_3 \log(age),$$
$$\tau = \beta_0 + \beta_1 age^{-2} + \beta_2 age^{-2} * \log(age) + \beta_3 age^{-1}).$$

Residual diagnostics did not indicate any significant violations of assumptions and the worm plot (Figure 4.45) and Q-statistics (Figure 4.46) indicated adequate model fit.

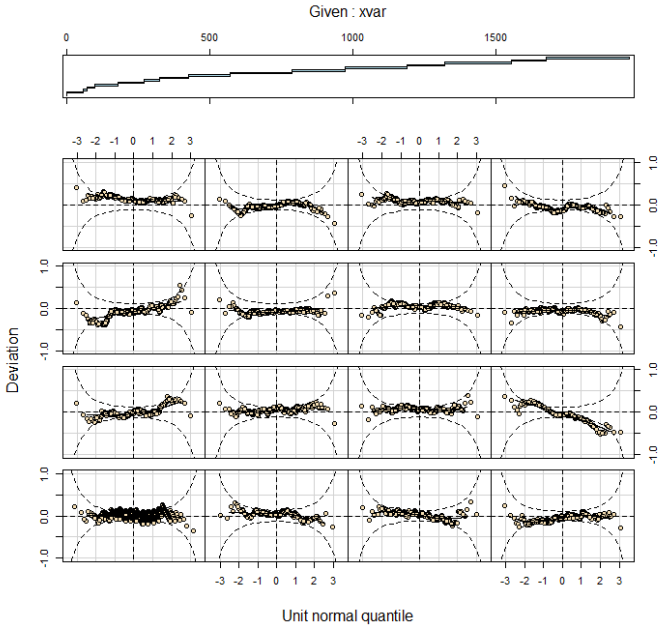


Figure 4.45: Worm plot for model using fractional polynomial smoothing for boys' height.

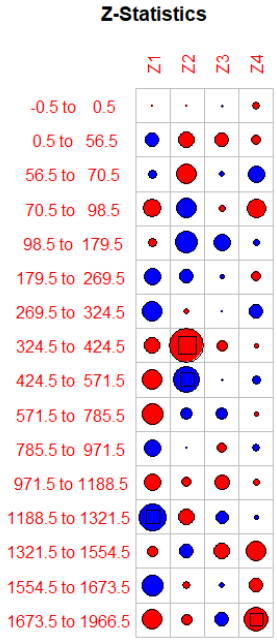


Figure 4.46: Q-statistics for model using fractional polynomial smoothing for boys' height.

The corresponding centile curves, shown in Figure 4.47 follow the pattern of the observed data well and achieve good agreement between the proportion of observations falling below each centile for the fitted model and the sample. They are nice and smooth over most of the age rage except for the 0.1 centile over the first few months.

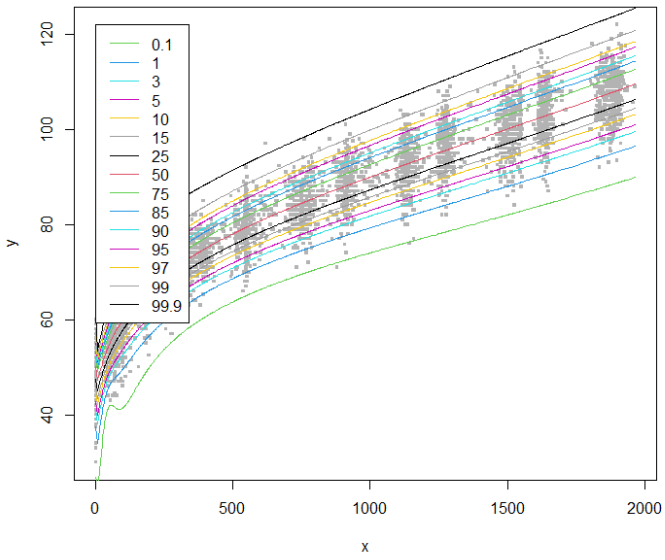
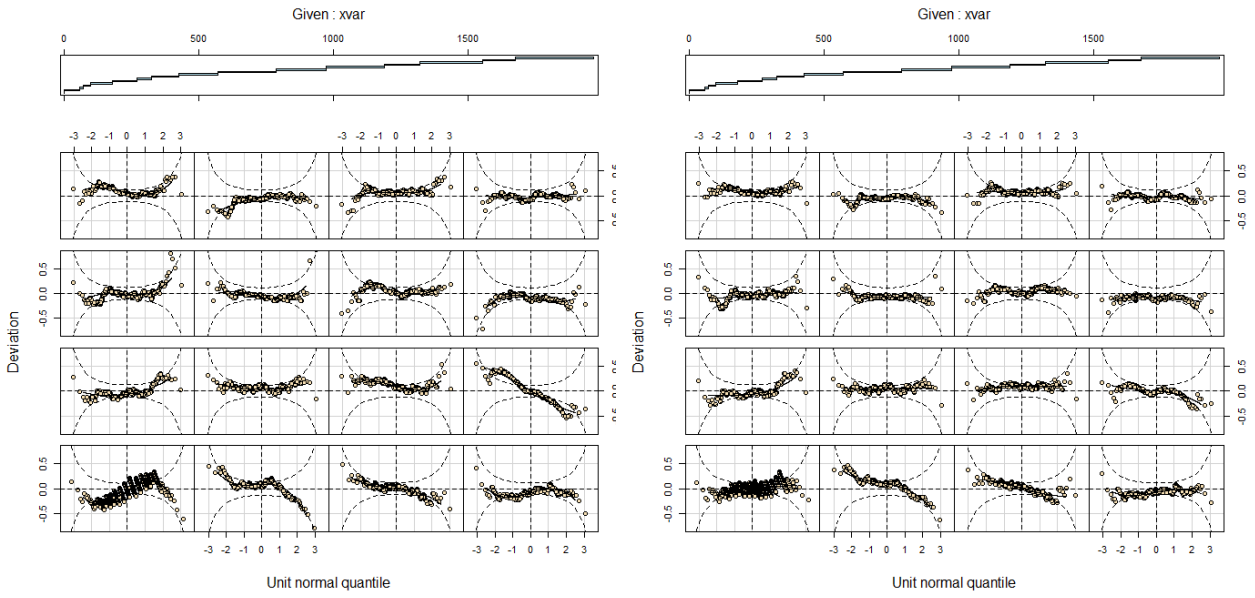


Figure 4.47: Centile curves model using fractional polynomial smoothing for boys' height.

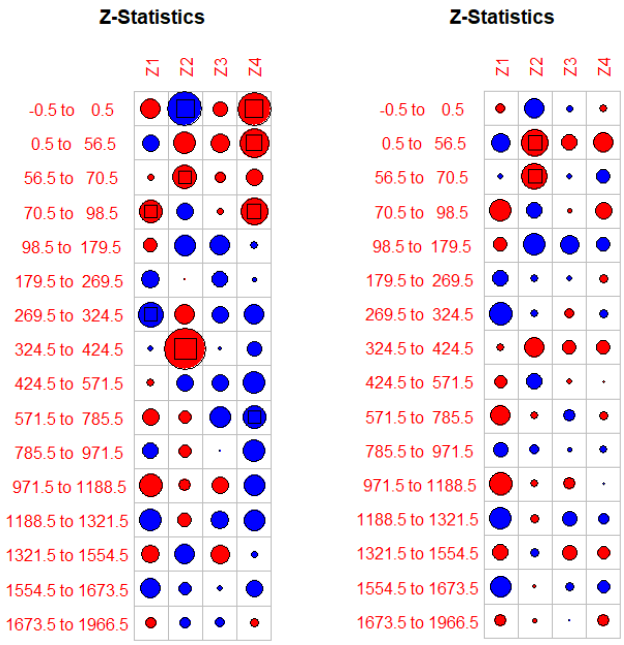
The Berkey-Reed model's assumptions of normality and constant variance were found to be more realistic for height. The BCT was the chosen distribution and the μ parameter was modelled using a Berkey-Reed First Order model. First the variance, skewness and kurtosis parameters were set to vary linearly with age. Thereafter, cubic smoothing splines were used to model the relationship of the parameters with age and the results were compared.

The model using cubic splines for the other three parameters produced a better fit as can be seen from the improvement in the worm plots (Figure 4.48) and Q-statistics (Figure 4.49), however the resulting curves for the more extreme centiles appear wavy, as can be seen in Figure 4.50. The centile curves produced when modelling σ, ν and τ linearly with age are nice and smooth and follow the pattern of the data well.



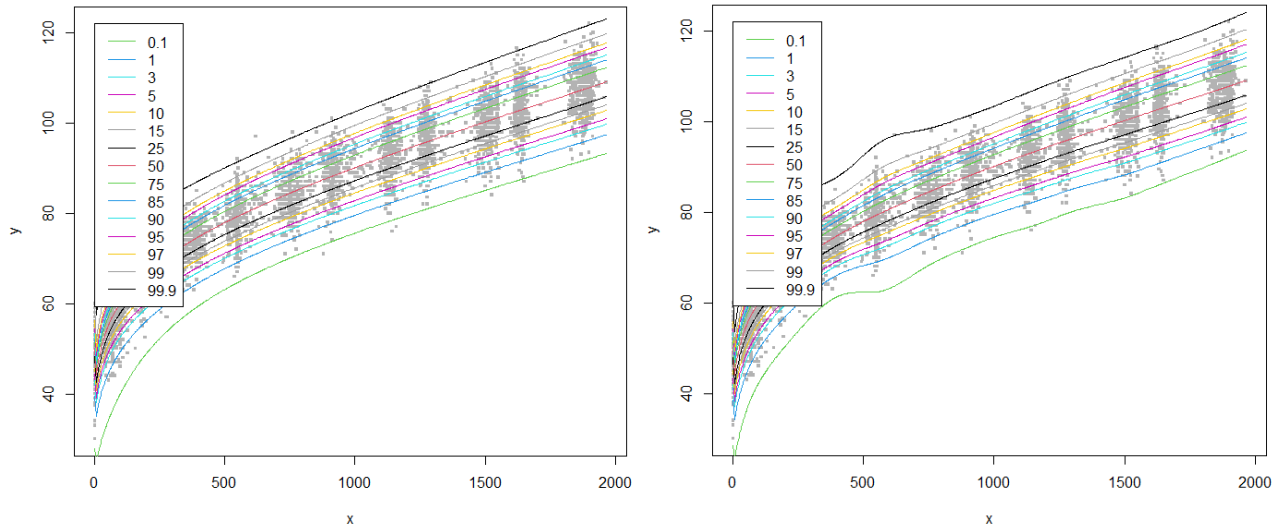
(a) **(b)**

Figure 4.48: Worm plots for BCT models using the Berkey-Reed First Order model for the μ parameter and (a) modelling the remaining parameters linearly with age or (b) using cubic splines.



(a) **(b)**

Figure 4.49: Q-statistics for BCT models using the Berkey-Reed First Order model for the μ parameter and (a) modelling the remaining parameters linearly with age or (b) using cubic splines.



(a) **(b)**

Figure 4.50: Centile curves for BCT models using the Berkey-Reed First Order model for the μ parameter and (a) modelling the remaining parameters linearly with age or (b) using cubic splines.

The Berkey-Reed Second Order model did not offer any improvement over the First Order Model. Based on a comparison of the GAIC(3) and AIC in Table 4.14, the First Order model with the remaining three parameters modelled as cubic splines with age is the preferred model. The more restrictive BIC criterion favours the First Order model with the remaining parameters modelled as linear functions of age.

Table 4.14: Berkey-Reed model comparison.

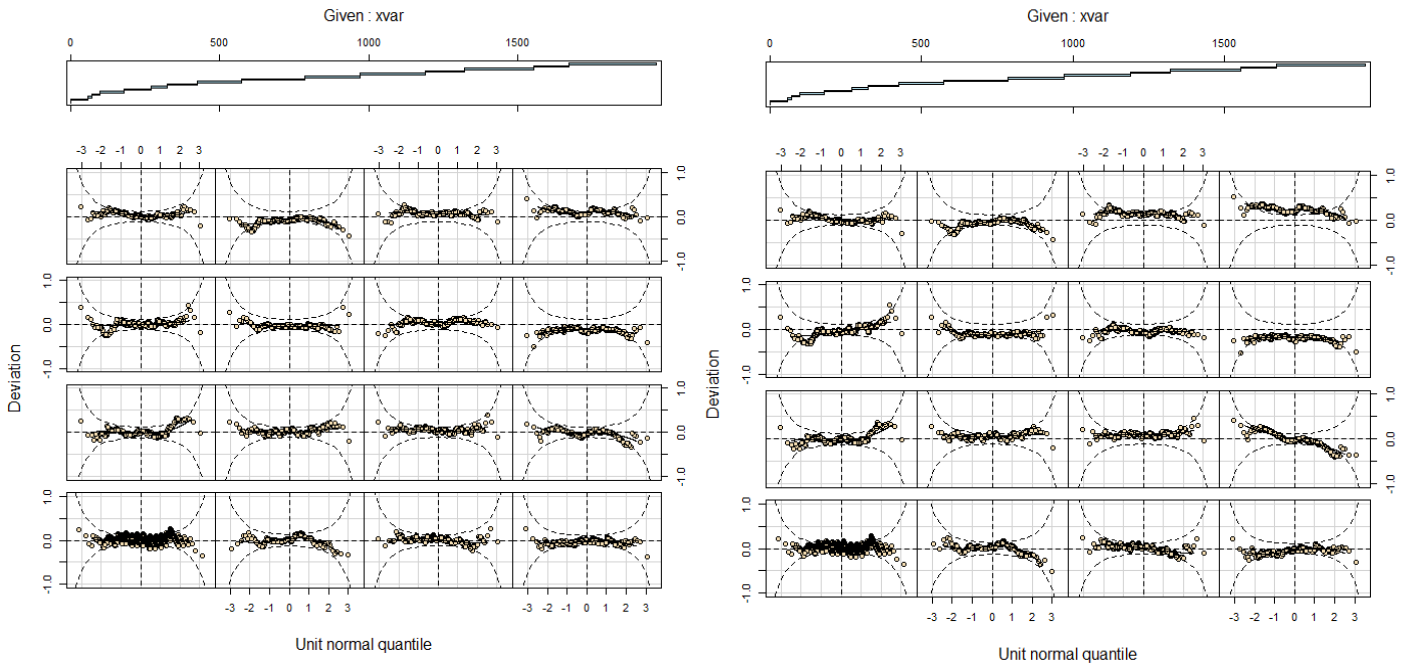
Model	**AIC**	**GAIC(3)**	**BIC**
First order, linear	39200.38	39210.38	**39268.93**
First order, spline	**39136.04**	**39162.04**	39314.28
Second order, linear	39202.38	39213.38	39277.79
Second order, spline	39138.15	39163.15	39309.54

Boys' heights were also modelled using the BCT distribution with curve smoothing by cubic splines with the model parameters selected automatically by the find.hyper() procedure. The final models selected based on penalties of *k=3* and *k=log(n)* are shown in Table 4.15.

Table 4.15: Comparing penalties for BCT model with cubic splines smoothing and automatic smoothing parameter selection.

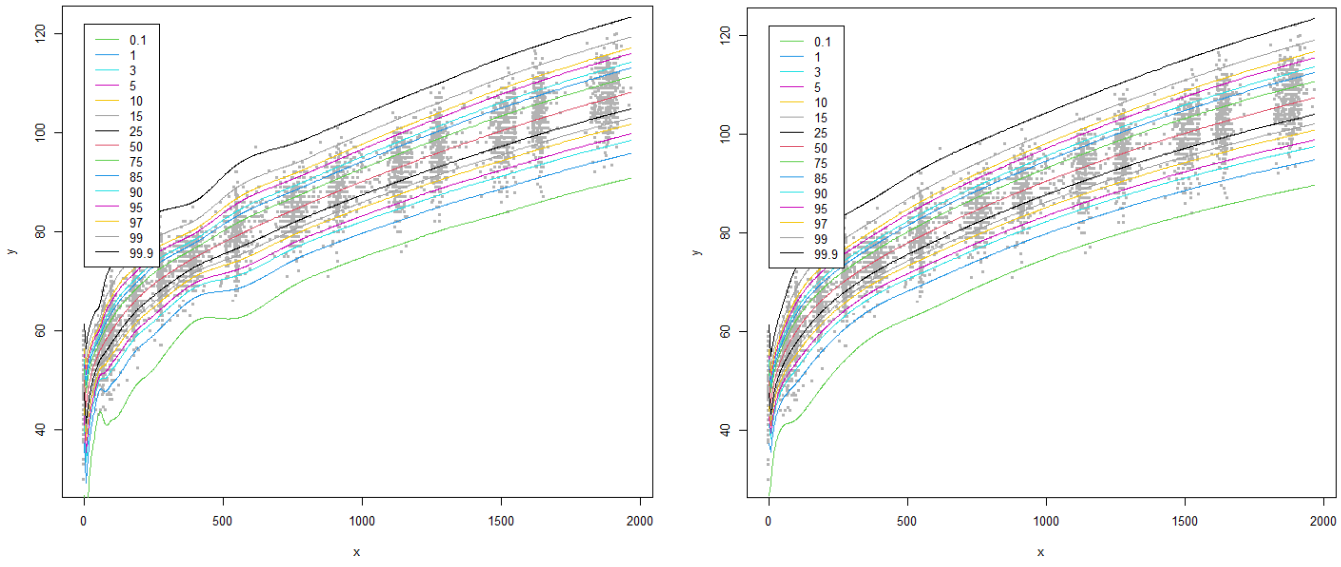
Penalty	**GAIC(3)**	**BIC**	df_{μ}	df_{σ}	df_{ν}	df_{τ}	ξ
BCT, *k*=3	39123.18	39310.12	5.7259	14.1179	1.4583	2.6201	0.0778
BCT, *k=log(n)*	39175.17	39299.71	3.7195	5.9305	1.4457	2.1722	0.0977

The worm plots and corresponding centile curves for the two models are plotted in Figures 4.51 and 4.52 respectively. The worm plots show that the model based on a penalty of *k=3* achieves a slightly better fit, however, the centile curves produced are unacceptably wavy. The more restrictive BIC criterion marginally sacrifices model fit in exchange for improved centile curves.



(a) **(b)**

Figure 4.51: Worm plots for BCT models with cubic spline smoothing and automatic smoothing parameter selection for (a) a penalty of $k=3$ and (b) a penalty of $k=log(n)$.



(a) **(b)**

Figure 4.52: Centile curves for BCT models with cubic spline smoothing and automatic smoothing parameter selection for (a) a penalty of $k=3$ and (b) a penalty of $k=log(n)$.

When implementing the WHO methodology for smoothing parameter selection, the BCPE distribution with curve smoothing by cubic splines was used. This was in line with the WHO's advice to keep the modelling approach the same across anthropometric measures in the interest of simplicity and consistency.

The final model was *BCPE(x=age^0.15,* df_μ*=4,* df_σ*=6,* df_ν*=2,* df_τ *=3)* which achieved adequate fit (Figure 4.53) and smooth centile curves (Figure 4.54) that follow the observed data well (Table 4.16).

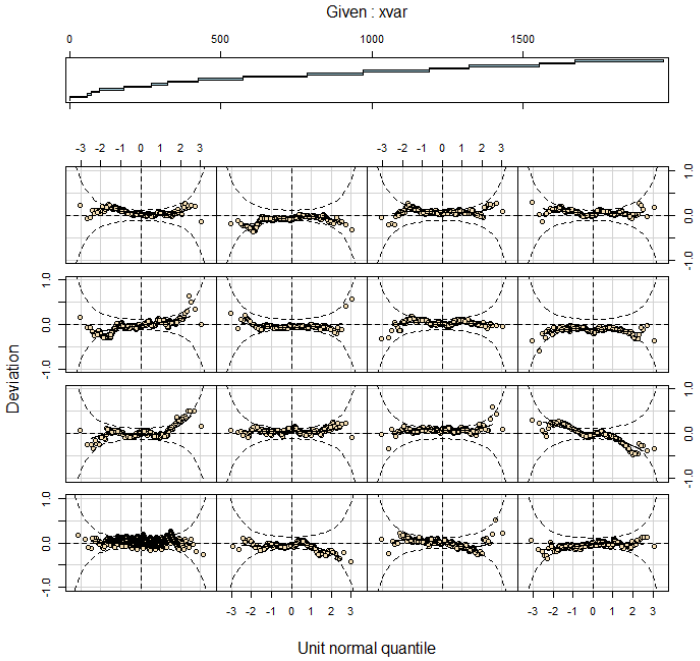


Figure 4.53: Worm plot for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

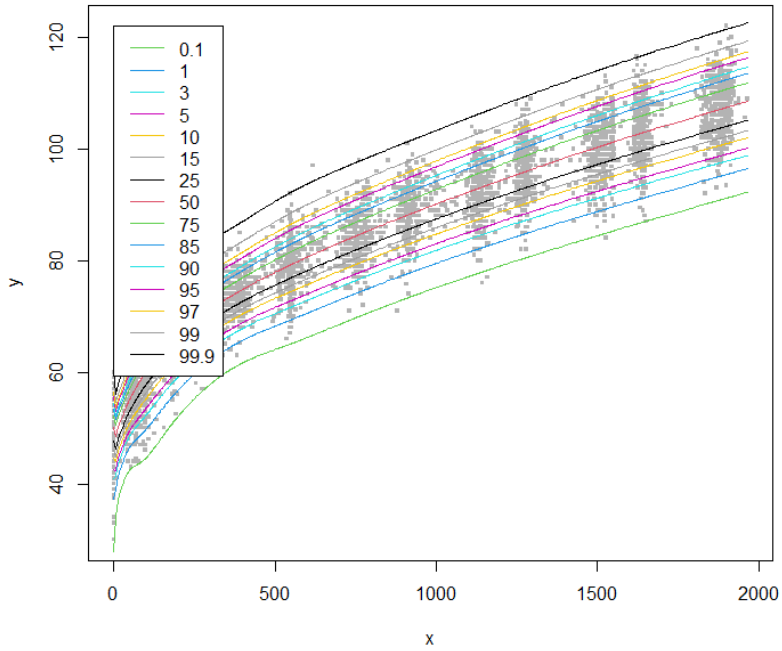


Figure 4.54: Centile curves for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

Table 4.16: Proportion of observations below each centile curve for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

Fitted centile	Proportion of observations falling below fitted centile
0.1	0.07130633
1	1.155163
3	3.137479
5	4.877353
10	9.512265
15	14.14718
25	25.08557
50	50.55619
75	75.51341
85	84.91158
90	90.34512
95	95.12265
97	97.26184

99	98.87336
99.9	99.84313

Table 4.17 below compares the parametric and semi-parametric models for boys' height based on the GAIC with various penalties. The final model selection was the Box-Cox Power Exponential distribution with curve smoothing by cubic splines and parameters determined following the WHO model selection procedure. This model achieved the lowest AIC, GAIC(3) as well as GAIC(6). The more restrictive BIC criterion favoured the simplicity of the model with the BCT distribution using the first order Berkey-Reed model for the μ parameter and linear relationships for the remaining parameters with age.

Table 4.17: Comparison of parametric and semi-parametric models for boys' height.

Model	**AIC**	**GAIC(3)**	**GAIC(6)**	**BIC**	**df**
BCT with cubic spline smoothing – automatic procedure	39153.90	39175.17	39238.98	39299.71	21.2705
BCPE with cubic spline smoothing – WHO procedure	**39119.54**	**39142.54**	**39211.54**	39277.21	23.0001
BCT with Berkey-Reed first order model for μ and linear relationships with age for σ, ν and τ	39200.38	39210.38	39246.38	**39268.93**	10.0000
BCT with fractional polynomials "smoothing"	39142.68	39170.68	39254.68	39334.63	28.0000

A quantile regression model with 7 degrees of freedom (knots at 10 weeks, 6 months, 18 months 48 months) was fitted to boys' heights to serve as an assessment of the choice of distribution for the chosen semi-parametric model. The worm plot in Figure 4.55 shows that the quantile regression model fits the data adequately. Figure 4.56 shows the quantile

regression curves overlaid on the semi-parametric curves and show close agreement between them except at the tail end of the 1st and 5th centile where the curves display some erratic behaviour. The close agreement among the rest of the curves was considered sufficient evidence that the final semi-parametric model selection fits the data well.

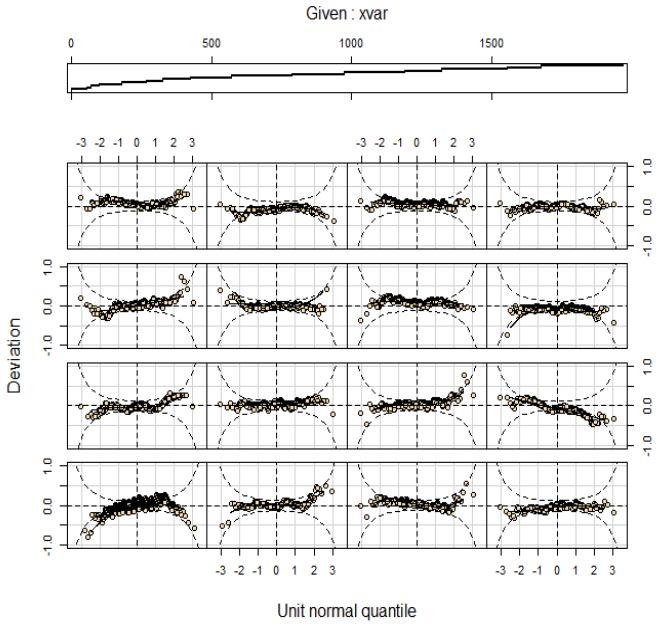


Figure 4.55: Worm plot for quantile regression model for boys' heights.

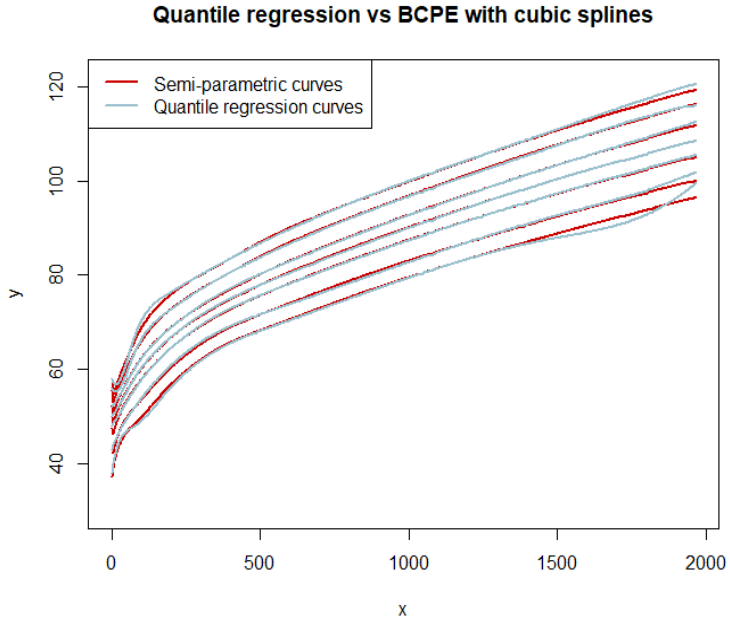


Figure 4.56: Comparison of centiles for quantile regression model and BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

4.4 Length/height-for-age: girls

The centile curves for girls' heights were derived using a semi-parametric model: the BCPE distribution and curve smoothing by cubic splines with the smoothing and transformation parameters determined according to the WHO methodology. The final model was *BCPE(x=age^0.15,* df_μ*=5,* df_σ*=5,* df_ν*=5,* df_τ *=1)*. The residual diagnostics were favourable and, as evidenced by the relatively horizontal worm plots in Figure 4.57 and the Q-statistics in Figure 4.58, the model fitted the data well.

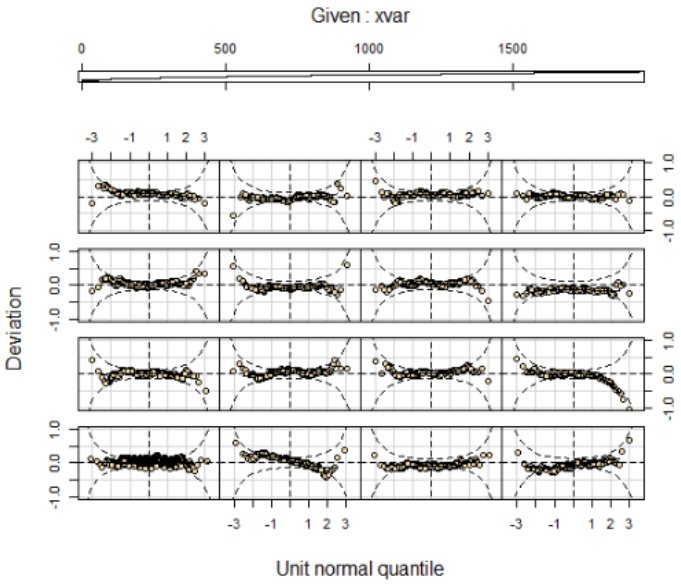


Figure 4.57: Worm plot for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

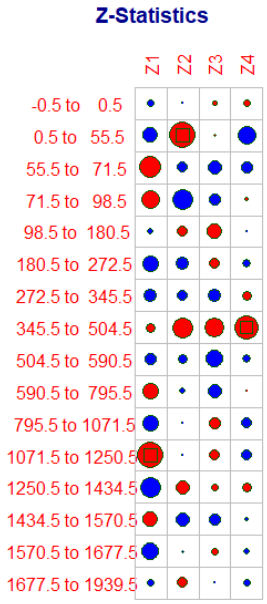


Figure 4.58: Q-statistics for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

The resulting centile curves are shown in Figure 4.59. The curves are nice and smooth over the age range with the exception of the 1st and 5th centiles which are a little wavy over the first few weeks.

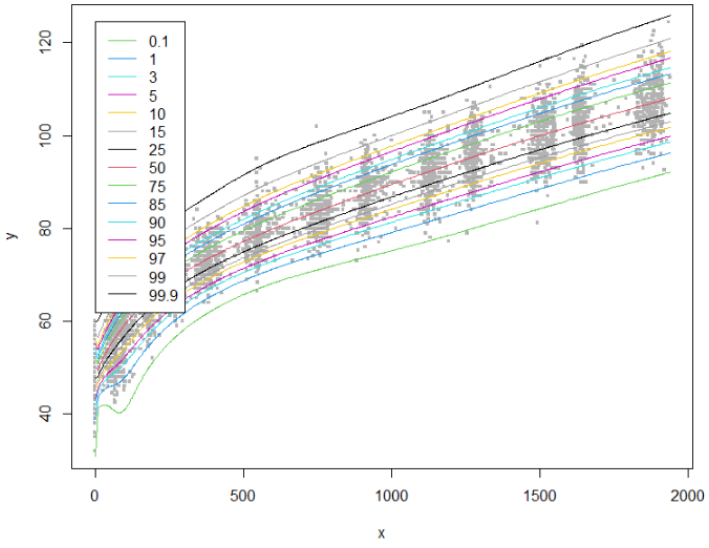


Figure 4.59: Centile curves for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

Figure 4.60 plots the centile curves of the semi-parametric model for girls' height with those of the quantile regression model overlaid. As previously done, knots were placed at 10 weeks, 6 months, 18 months, 48 months, resulting in a model with 7 degrees of freedom. The close agreement between the curves offered support for the appropriateness of the chosen semi-parametric model. In addition, the worm plot in Figure 4.61 confirms that the quantile regression model fits the data well.

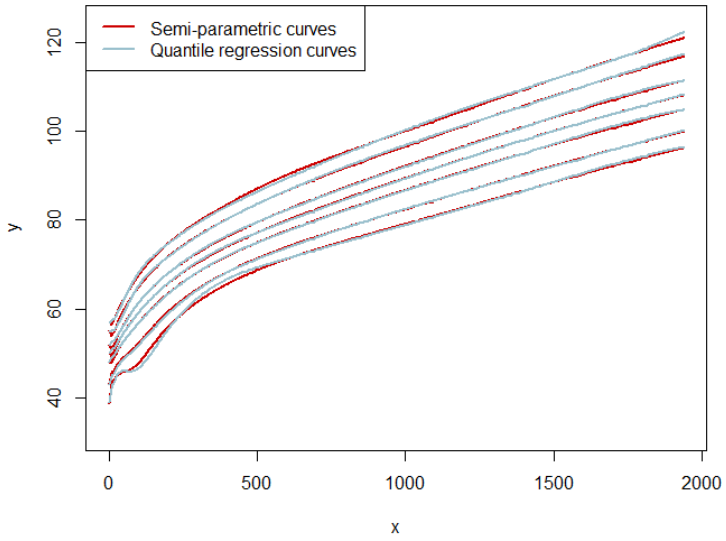


Figure 4.60: Centile curves of BCPE model with cubic spline smoothing and parameter selection according to WHO procedure with quantile regression model overlaid.

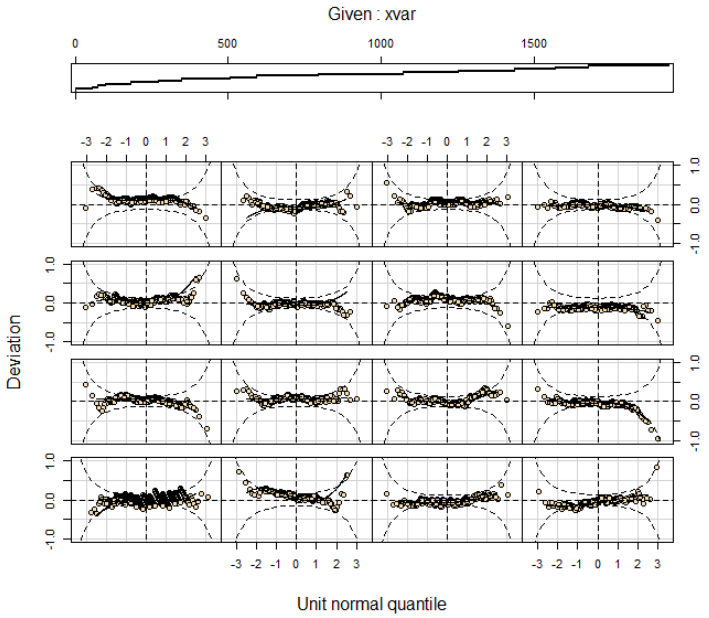


Figure 4.61: Worm plot for quantile regression model fitted to girls' height.

4.5 Head circumference-for-age: boys

As for the other growth measures, the starting point for deriving centile curves for head circumferences was to examine the density histograms with the theoretical densities of the Normal, BCCG, BCPE and BCT distributions overlaid. These curves suggested that the BCPE or BCT would be the most likely distribution selected.

The four distributions with different curve smoothing approaches were investigated and the best-fitting models are compared based on the GAIC in Table 4.18 below. The best-fitting model was the BCPE distribution with cubic spline smoothing where the smoothing parameters were determined using the automatic function in R. The final choice of model was the one which used the WHO procedure to determine the smoothing parameters as the GAIC values were only marginally higher and this was consistent with the approach used for the weight and height centiles. Similarly to the height data, the fractional polynomial and Berkey-Reed smoothing approaches were better suited to modelling head circumferences than they were to weight.

Table 4.18: Comparison of best-fitting parametric/semi-parametric models for boys' head circumference.

Model	AIC	GAIC(3)	GAIC(6)	BIC	df
BCPE with cubic spline smoothing – automatic procedure	**25827.39**	**25845.75**	**25900.85**	**25952.03**	**18.36442**
BCPE with cubic spline smoothing – WHO procedure	25830.68	25848.68	25902.69	25952.86	18.00177
BCPE with Berkey-Reed second order model for μ and linear relationships with age for σ, ν and τ	25924.66	25935.66	25968.66	25999.32	11
BCPE with fractional polynomials "smoothing"	25847.67	25867.67	25927.67	25983.40	20

All of the above models achieved adequate model fit and sufficiently smooth centile curves which modelled the pattern of growth in head circumference well. Figure 4.62 and Table 4.19 show that the 1st, 25th, 50th, 75th and 99th centiles for the competing models are very similar.

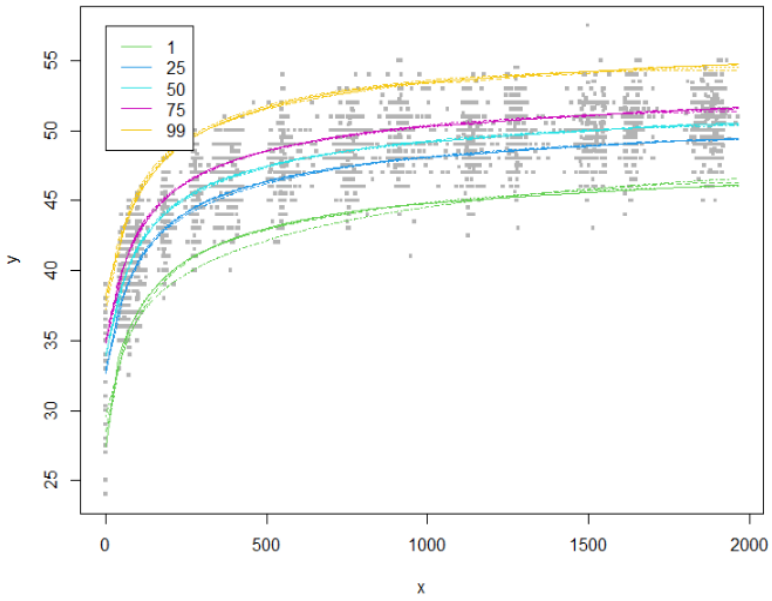


Figure 4.62: Centile curve comparison for competing parametric and semi-parametric models for boys' head circumference.

Table 4.19: Comparison of competing models for boys' head circumference based on percentage of cases below various centiles.

	% of cases below centile			
Centile	**BCPE with cubic spline smoothing – automatic procedure**	**BCPE with cubic spline smoothing – WHO procedure**	**BCPE with fractional polynomials "smoothing"**	**BCPE with Berkey-Reed second order model for μ and linear relationships with age for σ, ν and τ**
1	1.237019	1.16066	1.16066	1.053757
25	24.09896	23.74771	24.58766	25.80941
50	50.12217	49.66402	49.84728	50.39707
75	74.77092	75.91631	74.41967	74.64875
99	98.9157	99.00733	99.09896	99.12951

The residual diagnostics, worm plot and centile curves for the final chosen model, *BCPE(x=age^0.65,* df_μ*=7,* df_σ*=3,* df_ν*=0,* df_τ *=0)*, are shown in Figures 4.63 to 4.65 respectively. There is a random scatter in the plot of quantile residuals vs fitted values and there are no significant departures from the diagonal line in the Normal Q-Q plot. The worms in the worm plot are approximately horizontal indicating good fit of the model. The resulting centile curves are smooth and follow the pattern of the data well. In addition, the proportion of observations falling below each centile match the fitted centiles quite closely.

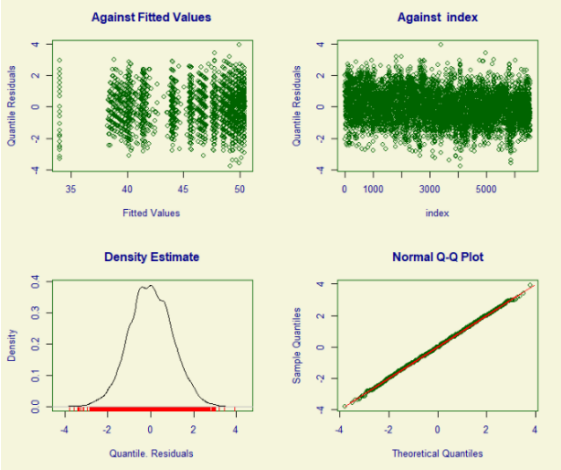


Figure 4.63: Residual diagnostics for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

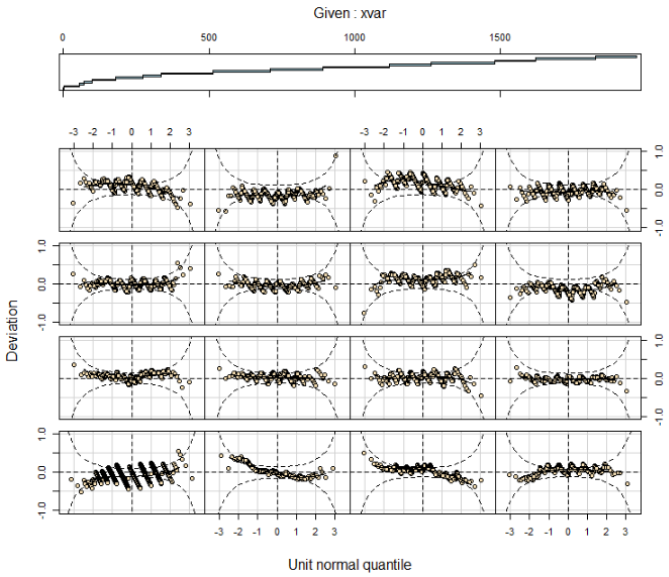


Figure 4.64: Worm plot for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

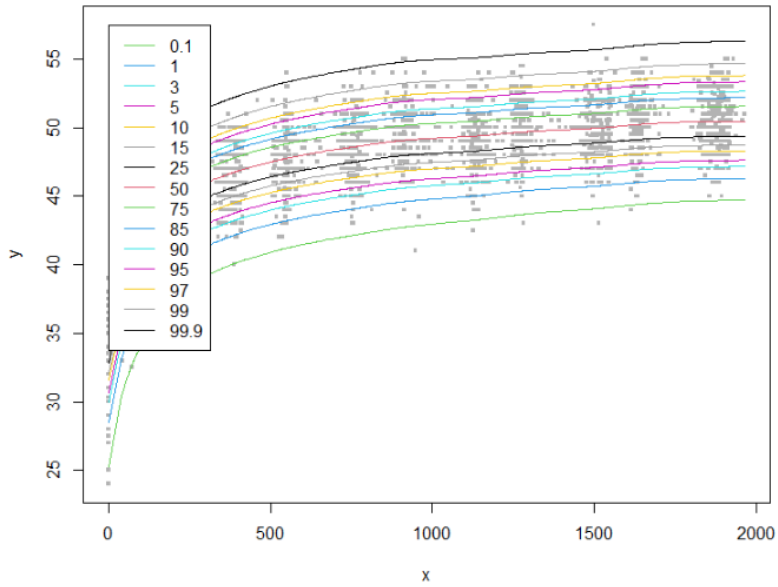


Figure 4.65: Centile curves for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

Table 4.20: Proportion of observations below each centile curve for final semi-parametric model for boys' head circumference.

Fitted centile	Proportion of observations falling below fitted centile
0.1	0.1527184
1	1.16066
3	2.962737
5	4.718998
10	9.499084
15	14.4777
25	23.74771
50	49.66402
75	75.91631
85	84.91142
90	90.33293
95	94.86866
97	96.7471
99	99.00733
99.9	99.93891

A quantile regression model, with knots at 10 weeks, 6 months, 18 months and 48 months, overlaid on the final semi-parametric model (Figure 4.66) for boys' head circumference further confirms the model selection. The worm plot in Figure 4.67 shows that the quantile regression model fits the data well.

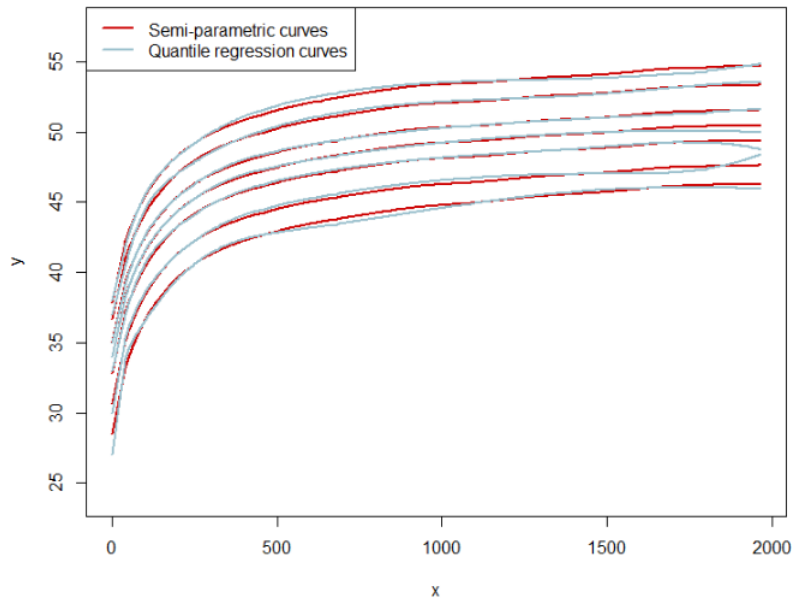


Figure 4.66: Centile curves of quantile regression model overlaid on BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

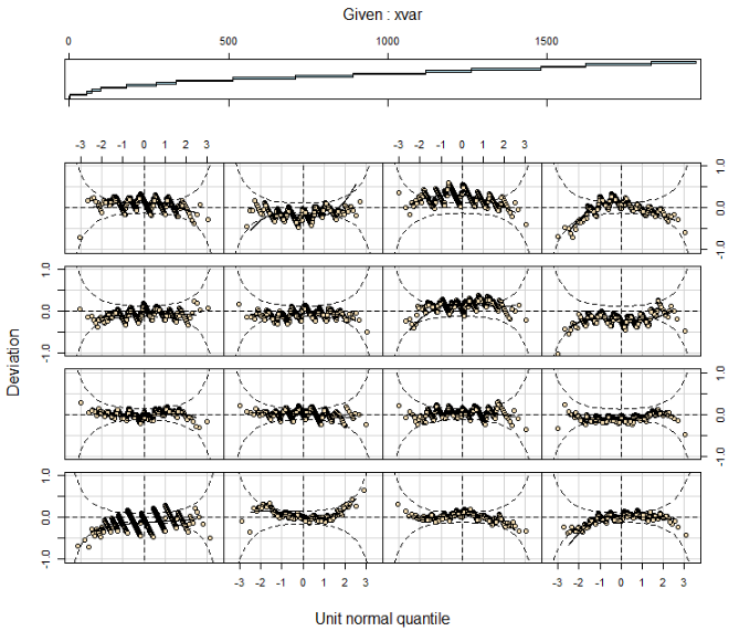


Figure 4.67: Worm plot for quantile regression model for boys' head circumference.

4.6 Head circumference-for-age: girls

The centile derivation for girls' head circumferences was done using the BCPE distribution with curve smoothing by cubic splines and the parameters determined according to the WHO procedure. The final model was *BCPE(x=age^0.65,* df_{μ}*=7,* df_{σ}*=2,* df_{ν}*=0,* df_{τ}*=0)* with the skewness and kurtosis parameters modelled linearly with age.

This model achieved reasonably good fit to the data, as evidenced by the worm plot in Figure 4.68, and produced smooth centile curves (Figure 4.69) with close agreement between the proportions of observed head circumference measurements below each centile curve and the expected proportions based on the fitted models (Table 4.21).

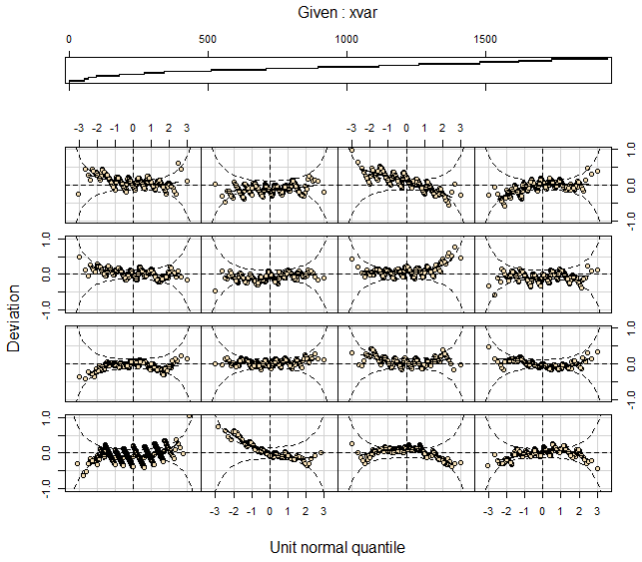


Figure 4.68: Worm plot for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

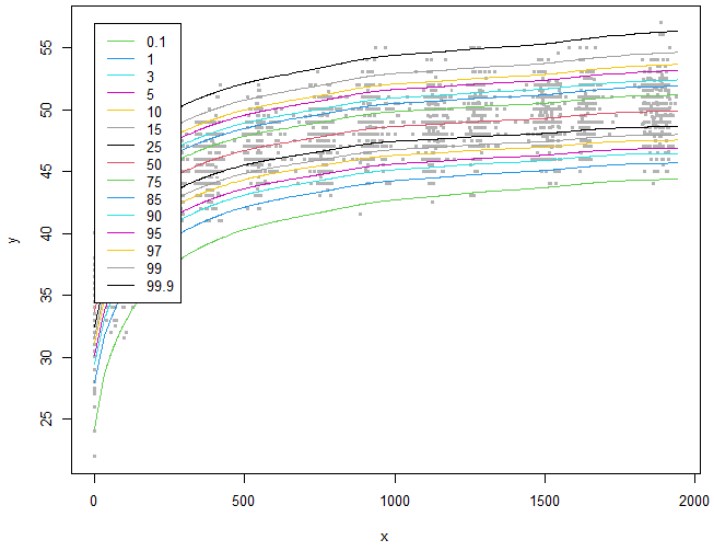


Figure 4.69: Centile curves for BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

Table 4.21: Proportion of observations below each centile curve for final semi-parametric model for girls' head circumference.

Fitted centile	Proportion of observations falling below fitted centile
0.1	0.1616031
1	1.195863
3	2.908856
5	4.395604
10	9.631545
15	14.09179
25	25.35553
50	48.91726
75	74.33743
85	84.50226
90	90.46542
95	95.57207
97	97.18811
99	98.85262
99.9	99.82224

Figure 4.70 shows that the quantile regression model for girls' head circumferences (with knots at 10 weeks, 6 months, 18 months and 48 months as for the other models) produces rather erratic curves, particularly towards the upper end of the age range. However, since the general pattern of growth is the same for the quantile regression curves and the semi-parametric curves, the former is considered sufficient evidence of fit of the latter to the data.

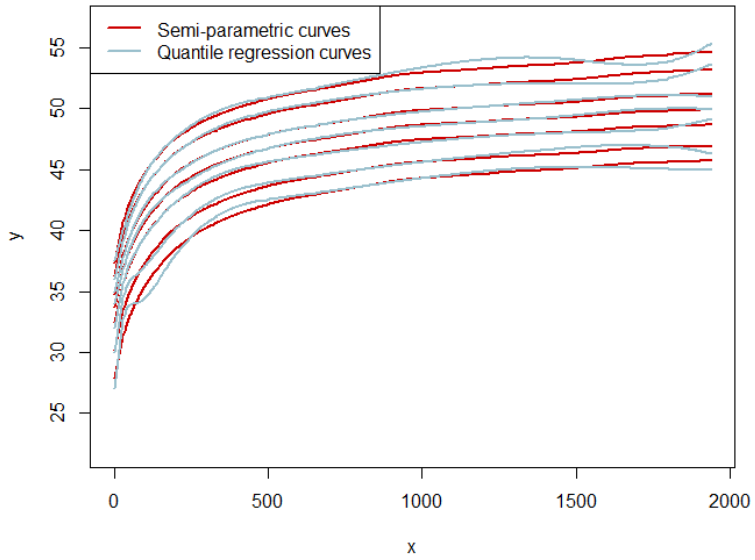


Figure 4.70: Quantile regression model curves overlaid on BCPE model with cubic spline smoothing and parameter selection according to WHO procedure.

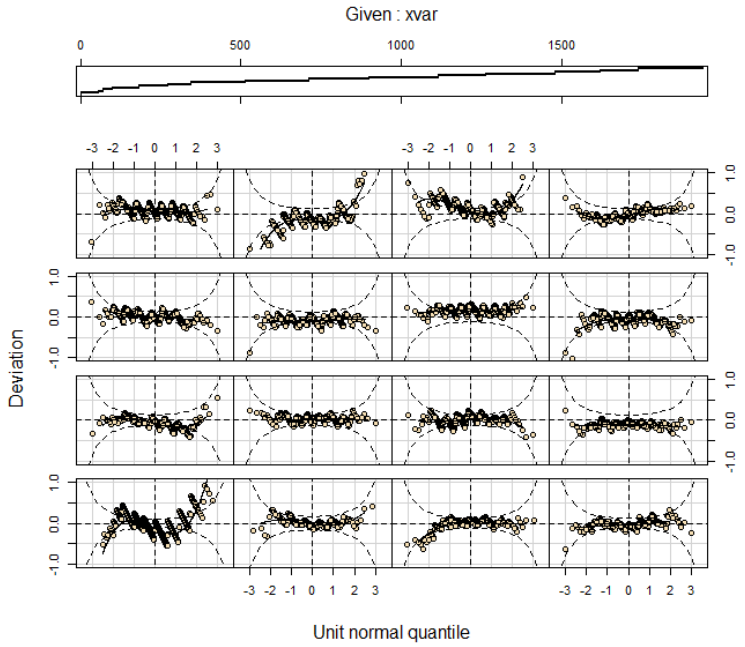


Figure 4.71: Worm plot for quantile regression model for girls' head circumference.

5 Comparison of derived centiles against WHO standards

5.1 Growth chart comparisons

A comparison was done between the growth references derived for the Drakenstein cohort and the WHO standards to serve as a further assessment of the growth of the children in relation to the standards. Figure 5.1 shows the derived centile curves with the WHO standards overlaid for boys' weight.

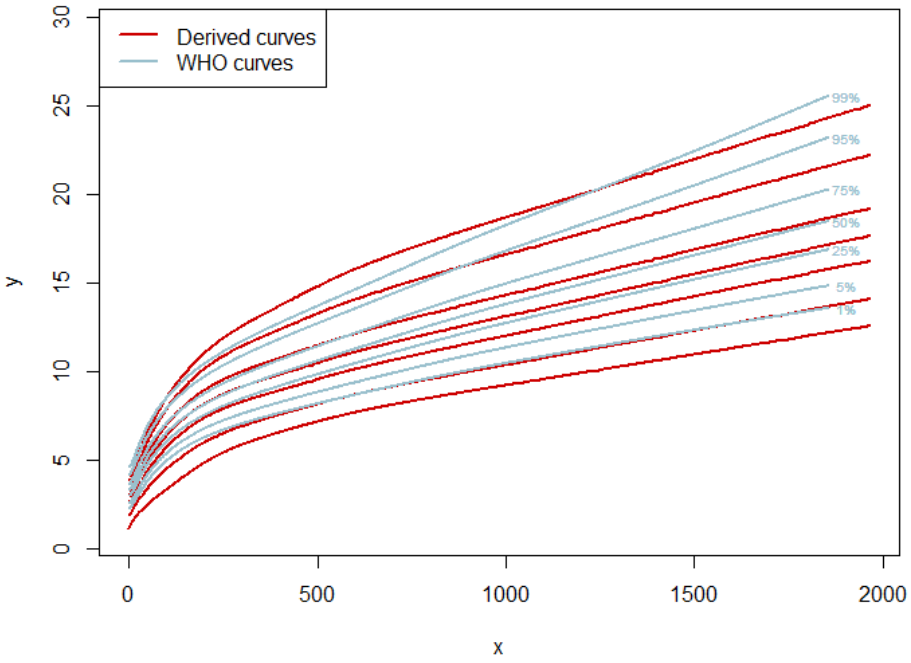


Figure 5.1: Comparison between growth references for Drakenstein boys' weights and WHO standards.

The 50th weight centile of the Drakenstein boys tracks closely with the WHO 50th centile until approximately 1.5 years of age, whereafter the Drakenstein boys lag behind. The same applies for the 25th centile, although the weights begin to lag behind those of the WHO even earlier – from approximately one year. There is a sustained significant difference between the growth curves for the lower centiles, with those of the Drakenstein boys lying below the WHO curves. For the curves above the 50th centile, the Drakenstein boys are either on par with or heavier than the WHO standards until approximately age two to three, after which the WHO boys overtake and are heavier at age 5.

For the lower centiles (1%, 5%, 25%, 50%, 75%), the Drakenstein girls' curves (Figure 5.2) lag behind those of the WHO subjects, with the gap widening with increasing age. For the 95th centile, however, the Drakenstein girls initially increase in weight more quickly than their WHO counterparts but they become relatively on par from approximately 3.5 years after which the WHO curve lies slightly above the Drakenstein one. For the most extreme centiles, 99% and beyond, the Drakenstein girls exceed the corresponding WHO curve quite substantially. Overall, the WHO centile curves for girls' weights span a narrower range of weights than the Drakenstein curves.

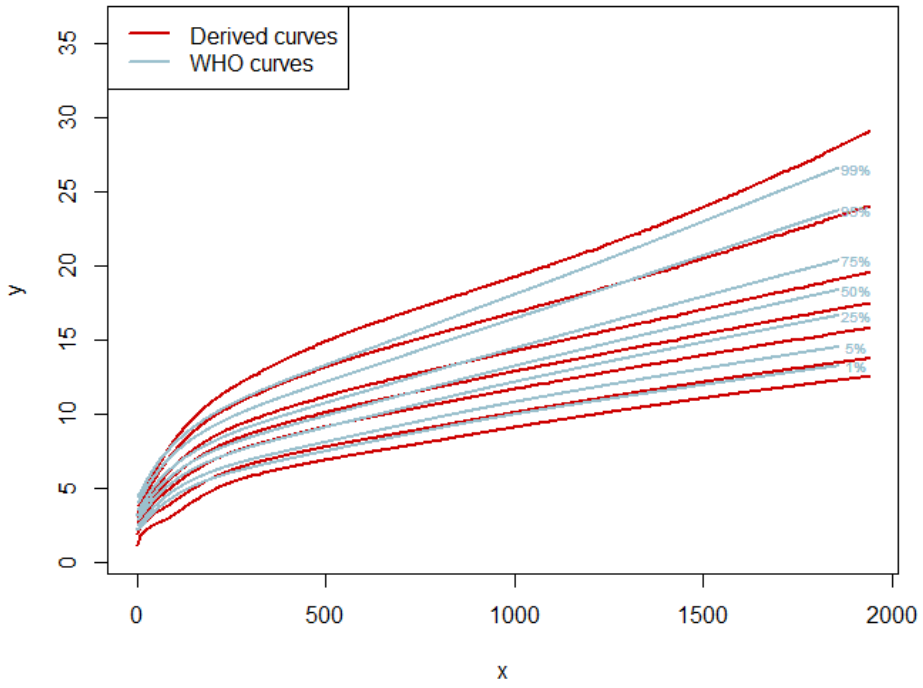


Figure 5.2: Comparison between growth references for Drakenstein girls' weights and WHO standards.

Figure 5.3 shows that the centile curves derived for Drakenstein boys' heights deviate quite substantially from the WHO curves - the only time frame where the curves are similar is approximately the first two months for the children in the 95th and 99th centiles for height. Thereafter, the Drakenstein curves illustrate a more rapid rate of growth until approximately age 12 months and 18 months for the 95th and 99th centiles respectively. Apart from these instances, the WHO curves all lie above the derived curves.

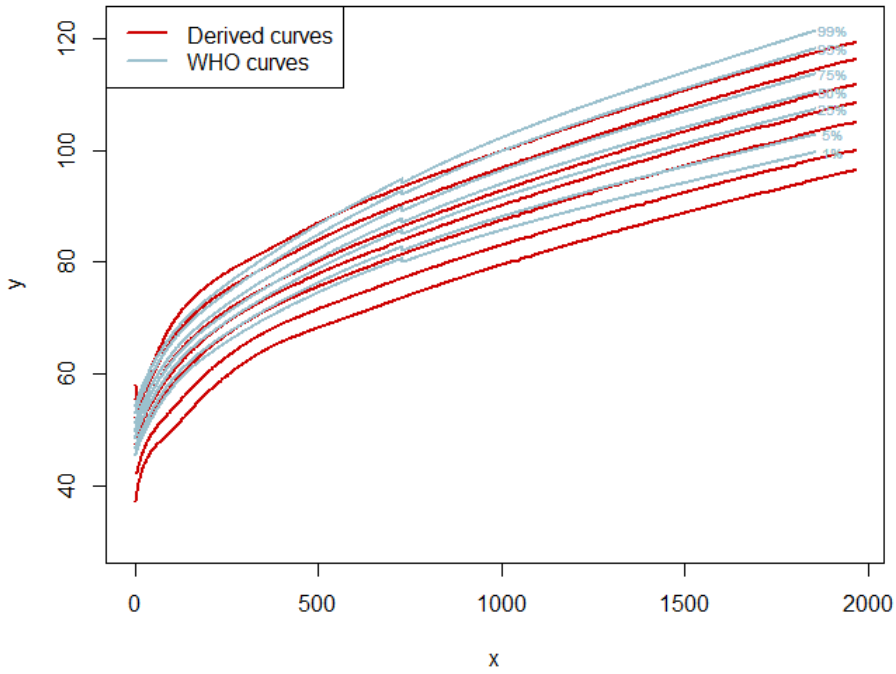


Figure 5.3: Comparison between growth references for Drakenstein boys' heights and WHO standards.

The differences observed between the derived curves for girls' heights and the WHO ones are very similar to those observed for boys' heights.

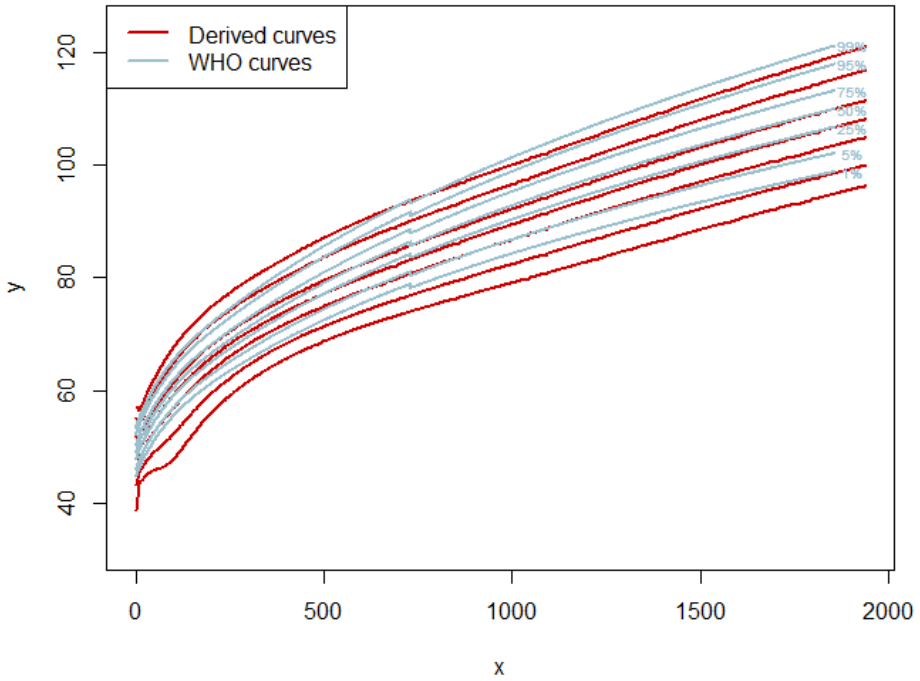


Figure 5.4: Comparison between growth references for Drakenstein girls' heights and WHO standards.

Figures 5.5 and 5.6 plot the derived Drakenstein growth references for boys' and girls' head circumferences respectively with the WHO standards overlaid. The upper extreme head circumference centiles (95th and 99th) of the Drakenstein boys exceed those of the WHO cohort. The opposite is seen for the lower extremes where the 1st and 5th centiles of the Drakenstein cohort's head circumferences lie below those of the WHO cohort. For the centiles between these extremes, the curves are similar. The distinction between the derived curves and WHO curves for girls' head circumferences is the same as was observed for the boys: the uppermost centiles are higher and the lower centiles are lower for the Drakenstein cohort compared to the WHO cohort.

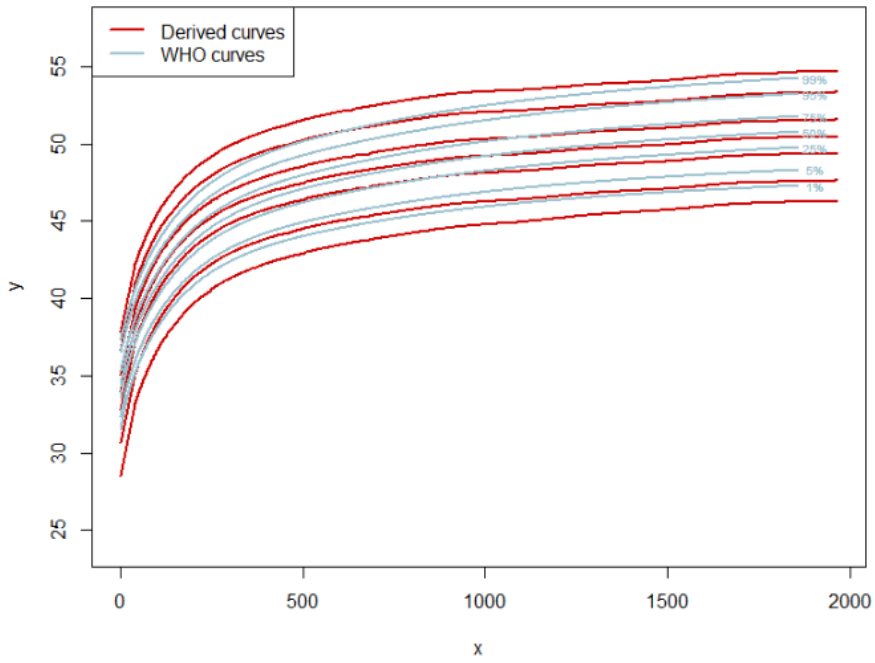


Figure 5.5: Comparison between growth references for Drakenstein boys' head circumferences and WHO standards.

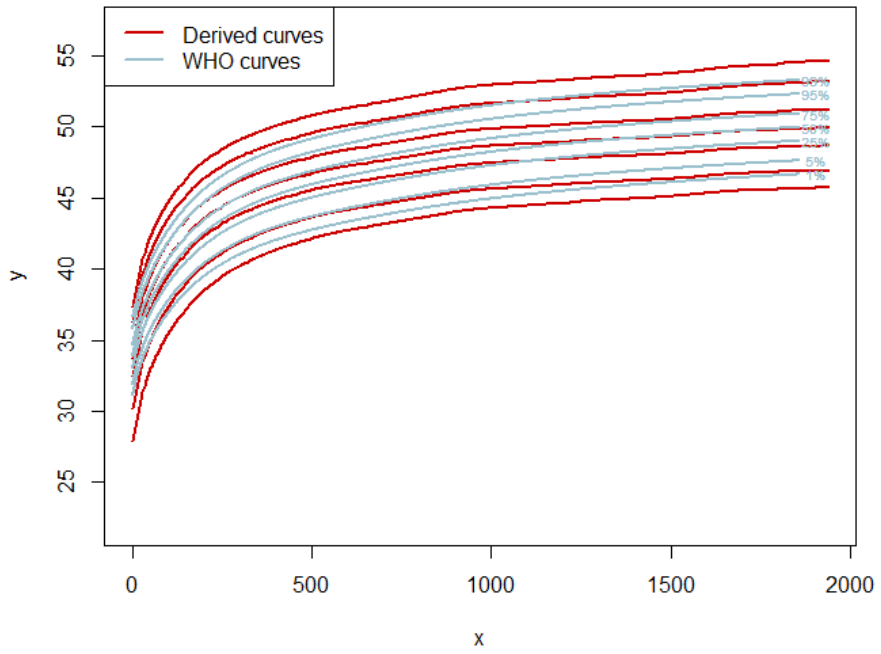


Figure 5.6: Comparison between growth references for Drakenstein girls' head circumferences and WHO standards.

5.2 Indicators of abnormal growth

Anthropometric indicators serve as a tool to identify abnormal growth patterns in children. The weight-for-age curves help to identify children who are underweight. The length/height for age curves may aid in diagnosing stunting (being abnormally short) as a result of undernutrition and/or recuring illness. In addition, the head circumference-for-age curves may highlight children with head circumferences that are too small (microcephaly) or too big (macrocephaly).

The aforementioned diagnoses are defined in terms of the z-scores for the growth measurements as follows:

- **Underweight**
  Underweight and severely underweight classifications correspond to weight-for-age z-scores below -2 and -3 respectively.
- **Stunted**
  Children are classified as stunted if they have a length/height-for-age z-score below -2 and severely stunted if the z-score is below -3.
- **Micro and macrocephaly**
  Micro- and macrocephaly refer to a having head circumference-for-age z-scores below -2 and above +2 respectively
  (WHO, 2008).

The derived growth references and WHO standards were used to classify growth measurements as normal or abnormal according to the above definitions. Comparisons were done across all visits as well as at the 24 months and 54 months visits. Table 5.1 below presents the classification of all boys' and girls' weight measurements for all visits as normal or underweight according to the WHO standards and in relation to deviation from the Drakenstein median in standard deviation units. For boys, of the 673 weights classified as underweight by the WHO, only 180 of them are underweight relative to the Drakenstein norm. For girls, 484 weights were classified as underweight by the WHO standards, but only 137 of those were also classified as underweight relative to the Drakenstein norm.

Table 5.1: Classification of weight measurements for all visits as normal or underweight based on WHO standards vs Drakenstein references.

Boys		**WHO standards**	
		Normal	Underweight
Drakenstein references	Normal	6187	**493**
	Underweight	0	180
Girls		**WHO standards**	
		Normal	Underweight
Drakenstein references	Normal	5843	**347**
	Underweight	0	137

Tables 5.2 and 5.3 present the classifications for weights as at the 24 month and 54 month visits respectively. At 24 months there were 25 boys and 18 girls classified as underweight despite normal classifications according to the Drakenstein references. At 54 months, there were 19 boys and 12 girls with such misclassifications.

Table 5.2: Classification of weight measurements as at the 24 month visit as normal or underweight based on WHO standards vs Drakenstein references.

Boys		**WHO standards**	
		Normal	Underweight
Drakenstein references	Normal	346	**25**
	Underweight	0	10
Girls		**WHO standards**	
		Normal	Underweight
Drakenstein references	Normal	336	**18**
	Underweight	0	8

Table 5.3: Classification of weight measurements as at the 54 month visit as normal or underweight based on WHO standards vs Drakenstein references.

Boys		**WHO standards**	
		Normal	Underweight
Drakenstein references	Normal	416	**19**
	Underweight	0	9
Girls		**WHO standards**	
		Normal	Underweight
Drakenstein references	Normal	395	**12**
	Underweight	0	9

Similar findings were made for heights and head circumferences. The stunting classifications for all visits as well as at the 24 month and 54 month visits appear in Tables 5.4-5.6. At 24 months, 79 boys and 48 girls were classified as stunted based on the WHO standards but not the Drakenstein references. At 54 months, the numbers were 40 and 34 for boys and girls respectively. There were also a number of children classified as having micro- or macrocephaly based on the WHO standards but not according to the Drakenstein references. These numbers are highlighted in Tables 5.7-5.9.

Table 5.4: Classification of height measurements for all visits as normal or stunted based on WHO standards vs Drakenstein references.

Boys		**WHO standards**	
		Normal	Stunted
Drakenstein references	Normal	5288	**1232**
	Stunted	0	166
Girls		**WHO standards**	
		Normal	Stunted
Drakenstein references	Normal	5267	**784**
	Stunted	0	147

Table 5.5: Classification of height measurements as at the 24 months visit as normal or stunted based on WHO standards vs Drakenstein references.

Boys		**WHO standards**	
		Normal	Stunted
Drakenstein references	Normal	293	**79**
	Stunted	0	8
Girls		**WHO standards**	
		Normal	Stunted
Drakenstein references	Normal	302	**48**
	Stunted	0	11

Table 5.6: Classification of height measurements as at the 54 months visit as normal or stunted based on WHO standards vs Drakenstein references.

Boys		**WHO standards**	
		Normal	Stunted
Drakenstein references	Normal	395	**40**
	Stunted	0	7
Girls		**WHO standards**	
		Normal	Stunted
Drakenstein references	Normal	368	**34**
	Stunted	0	12

Table 5.7: Classification of head circumference measurements for all visits as normal, microcephaly or macrocephaly based on WHO standards vs Drakenstein references.

Boys		**WHO standards**		
		Normal	Microcephaly	Macrocephaly
Drakenstein references	Normal	5390	**193**	**355**
	Microcephaly	0	143	0
	Macrocephaly	0	0	142
Girls		**WHO standards**		
		Normal	Microcephaly	Macrocephaly
Drakenstein references	Normal	4961	**170**	**481**
	Microcephaly	0	146	0
	Macrocephaly	0	0	139

Table 5.8: Classification of head circumference measurements as at the 24 months visit as normal, microcephaly or macrocephaly based on WHO standards vs Drakenstein references.

Boys		**WHO standards**		
		Normal	Microcephaly	Macrocephaly
Drakenstein references	Normal	338	**11**	**18**
	Microcephaly	0	6	0
	Macrocephaly	0	0	6
Girls		**WHO standards**		
		Normal	Microcephaly	Macrocephaly
Drakenstein references	Normal	321	**2**	**18**
	Microcephaly	0	14	0
	Macrocephaly	0	0	6

Table 5.9: Classification of head circumference measurements as at the 54 months visit as normal, microcephaly or macrocephaly based on WHO standards vs Drakenstein references.

Boys		**WHO standards**		
		Normal	Microcephaly	Macrocephaly
Drakenstein references	Normal	412	**7**	**2**
	Microcephaly	0	9	0
	Macrocephaly	0	0	13
Girls		**WHO standards**		
		Normal	Microcephaly	Macrocephaly
Drakenstein references	Normal	385	**1**	**14**
	Microcephaly	0	10	0
	Macrocephaly	0	0	5

The above classification tables and overlaid centile plots suggest that the Drakenstein cohort as a whole differed from the WHO expected growth in terms of weight, height and head circumference. This could be attributed to the poor socioeconomic status and harmful exposures of the Drakenstein cohort which have been shown to negatively impact childhood growth. It is however also possible that South African children as a whole grow differently to the growth patterns represented by the WHO standards. Conducting a fair comparison would require a reasonably healthy and representative sample of the South African population.

6 Conclusions

Childhood growth monitoring with growth curves which accurately represent the growth of the population of interest is paramount to ensuring that abnormal growth is identified and the appropriate interventions implemented timeously. In light of numerous studies which have found certain population groups to deviate from the widely accepted WHO (2006) Child Growth Standards, this dissertation sought to model the growth of a cohort of South African children and compare their growth to these standards.

Initial visual assessments of the observed anthropometry measures overlaid on the WHO standards as well as the WHO z-scores calculated for the Drakenstein cohort suggested the presence of deviations in their patterns of growth compared to the WHO cohort.

Different approaches to dealing with longitudinal data in the derivation of growth centile curves were discussed and the decision was made to ignore the serial nature of the data and proceed as though the data were cross-sectional.

Growth reference curves were derived for the Drakenstein cohort using parametric, semi-parametric and non-parametric methods and these were compared based on the model fit and appearance of the resultant centile curves. Model fit was assessed in terms of the GAIC, worm plots and Q-tests. The parametric and semi-parametric methods were implemented within the GAMLSS framework and the Normal, Box-Cox Cole and Green, Box-Cox Power Exponential and the Box-Cox T distributions were investigated. Curve smoothing was done by cubic splines, fractional polynomials and the Berkey-Reed First and Second Order models.

The methods considered were chosen based on their ability to model the rapid growth in early infancy followed by decelerated growth into childhood, as well as deal with the presence of skewness and kurtosis in the data. The preferred approach for all of the growth measures assessed – weight, height and head circumference – was to use the Box-Cox Power Exponential distribution and curve smoothing by cubic splines with the smoothing parameters determined based on the methodology used by the WHO when deriving the child growth standards. The fractional polynomial smoothing and the use of the Berkey-Reed

model for the location parameter did not fare particularly well for the weight data but performed better when modelling height and head circumference.

Non-parametric quantile regression, while not recommended as a primary approach for deriving growth centiles proved valuable in assessing the appropriateness and fit of the final chosen parametric distributions.

A method for calculating confidence intervals for the centile curves using bootstrapping was discussed and demonstrated for boys' weights.

The derived Drakenstein growth references and the WHO standards were compared visually as well as through the comparison of classifications of abnormal growth. The occurrence of diagnoses of underweight-for-age, stunted and micro- and macrocephaly were cross-tabulated for both systems for all measurements as well as at 24 months and 54 months. It was found that measuring the Drakenstein cohort against the WHO standards resulted in a greater number of diagnoses of all the abnormal growth measures assessed.

This calls into question the appropriateness of the WHO standards for measuring South African children and emphasises the potential value of deriving national growth standards.

A caveat was however noted about this research: the cohort used is from a particularly disadvantaged environment which negatively impacts health and thereby also growth and development. It is possible that the differences between the WHO standards and the SA references are exaggerated because of this. In order to produce reliable national growth curves, it would be necessary to draw reasonably healthy study participants that were representative of the overall population that would serve as a standard for the growth that should be achieved by the children of South Africa.

Appendix A

Table A.1: Number of weight measurements per visit.

Visit	Boys	Girls	Total
Birth	579	550	1129
6-10 weeks	468	426	894
10 weeks	464	434	898
14 weeks	482	451	933
6 months	477	452	929
9 months	435	404	839
12 months	397	386	783
18 months	423	401	824
24 months	381	362	743
30 months	381	360	741
36 months	392	379	771
42 months	433	412	845
48 months	419	402	821
54 months	444	416	860
60 months	422	390	812
Other*	589	395	984

*Other encompasses all ad hoc visits associated with pneumonia episodes.

Table A.2: Number of height measurements per visit.

Visit	Boys	Girls	Total
Birth	571	539	1110
6-10 weeks	457	419	876
10 weeks	446	415	861
14 weeks	467	433	900
6 months	466	446	912
9 months	424	398	822
12 months	389	380	769
18 months	423	401	824
24 months	380	361	741
30 months	378	357	735
36 months	392	376	768
42 months	431	412	843
48 months	418	401	819
54 months	442	414	856
60 months	422	390	812
Other*	506	349	855

*Other encompasses all ad hoc visits associated with pneumonia episodes.

Table A.3: Number of head circumference measurements per visit.

Visit	Boys	Girls	Total
Birth	573	547	1120
6-10 weeks	459	417	876
10 weeks	454	422	876
14 weeks	473	444	917
6 months	474	453	927
9 months	429	403	832
12 months	393	386	779
18 months	422	401	823
24 months	379	361	740
30 months	380	362	742
36 months	392	377	769
42 months	437	414	851
48 months	419	399	818
54 months	443	415	858
60 months	421	387	808

Table A.4: Weight (in kg) summary by visit for boys and girls.

Visit	Boys (median[Q1-Q3])	Girls (median[Q1-Q3])
Birth	3.110 [2.685-3.440]	3.080 [2.730-3.400]
6-10 weeks	5.000 [4.4675-5.5800]	4.680 [4.220-5.180]
10 weeks	5.600 [5.0400-6.1650]	5.255 [4.600-5.7875]
14 weeks	6.395 [5.700-6.980]	5.900 [5.250-6.520]
6 months	8.100 [7.300-8.7400]	7.460 [6.700-8.230]
9 months	9.100 [8.225-10.005]	8.590 [7.6925-9.400]
12 months	9.520 [8.760-10.460]	9.250 [8.245-10.1875]
18 months	10.900 [9.980-12.000]	10.530 [9.420-11.700]
24 months	11.900 [10.800-12.800]	11.400 [10.325-12.500]
30 months	12.900 [11.800-14.000]	12.795 [11.500-14.000]
36 months	13.500 [12.600-14.700]	13.500 [12.200-14.750]
42 months	14.600 [13.500-15.900]	14.350 [13.000-15.800]
48 months	15.300 [14.200-16.600]	15.200 [13.800-17.075]
54 months	16.300 [15.075-17.900]	16.200 [14.800-18.000]
60 months	17.150 [15.825-19.000]	17.200 [15.500-19.300]

Table A.5: Height (in cm) summary by visit for boys and girls.

Visit	Boys (median[Q1-Q3])	Girls (median[Q1-Q3])
Birth	50.0 [48.00-52.00]	50.0 [48.00-52.00]
6-10 weeks	56.0 [53.50-58.00]	55.0 [53.00-56.00]
10 weeks	57.5 [55.00-59.875]	57.0 [54.25-59.00]
14 weeks	60.0 [57.50-62.00]	59.5 [57.00-61.00]
6 months	66.5 [64.00-68.00]	65.0 [63.00-67.00]
9 months	71.0 [69.00-73.00]	69.5 [67.50-72.00]
12 months	74.0 [72.00-76.00]	74.0 [71.00-76.00]
18 months	79.0 [76.95-82.00]	78.5 [76.00-81.50]
24 months	85.0 [82.00-87.00]	83.5 [81.00-86.00]
30 months	88.0 [86.00-91.00]	88.0 [85.00-90.50]
36 months	92.0 [89.80-95.00]	92.0 [89.00–95.00]
42 months	96.0 [93.35-99.00]	96.0 [93.00-98.50]
48 months	100.0 [97.00-103.50]	99.8 [97.00-103.00]
54 months	103.0 [100.00-106.175]	103.0 [100.00-106.00]
60 months	107.0 [104.00-110.975]	107.0 [104.00-110.00]

Table A.6: Head circumference (in cm) summary by visit for boys and girls.

Visit	Boys (median[Q1-Q3])	Girls (median[Q1-Q3])
Birth	34.0 [33.00-35.00]	34.0 [32.0-35.0]
6-10 weeks	39.0 [38.00-40.00]	38.0 [37.0-39.5]
10 weeks	40.0 [39.00-41.00]	40.0 [38.5-41.0]
14 weeks	41.6 [41.00-43.00]	41.0 [39.5-42.0]
6 months	44.0 [43.00-45.00]	43.0 [42.0-44.5]
9 months	46.0 [45.00-47.00]	45.0 [43.5-46.0]
12 months	47.0 [45.50-48.00]	46.0 [45.0-47.0]
18 months	48.0 [47.00-49.00]	47.0 [46.0-48.0]
24 months	48.5 [47.00-49.50]	48.0 [46.0-49.0]
30 months	49.0 [48.00-50.125]	49.0 [47.5-50.0]
36 months	49.0 [48.00-50.00]	48.5 [47.0-50.0]
42 months	50.0 [49.00-51.00]	49.0 [48.0-50.5]
48 months	50.0 [49.00-51.00]	49.0 [48.0-50.0]
54 months	50.5 [49.50-51.50]	50.0 [49.0-51.0]
60 months	50.0 [49.00-51.50]	50.0 [48.8-51.0]

List of References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716-723.

Altman, D. and Chitty, L. (1994). Charts of fetal size: 1 Methodology. *British Journal of Obstetrics and Gynaecology*, 101:29-34.

Berkey, C. and Reed, R. (1987). A Model for Describing Normal and Abnormal Growth in Early Childhood. *Human Biology*, 59 No. 6: 973-987.

Borghi, E., de Onis, M., Garza, C., Van den Broeck, J., Frongillo, E., Grummer-Strawn, L., Van Buuren, S., Pan, H., Molinari, L., Martorell, R., Onyango, A., Martines, J. (2006). Construction of the World Health Organisation child growth standards: selection of methods for attained growth curves. *Statistics in Medicine*, 25:247-265.

Cole, T. and Green, P. (1992). Smoothing reference centile curves: the LMS method and penalised likelihood. *Statistics in Medicine*, 11(10):1305-1319.

Cole, T.J. (2012). The development of growth references and growth charts. *Annals of human biology*, 39(5), 382–394.

Cole, T. (2019). Commentary: Methods for calculating growth trajectories and constructing growth centiles. *Statistics in Medicine*, 38: 3571-3579.

Cole, T. (2021). Sample size and sample composition for constructing growth reference centiles. *Statistical Methods in Medical Research*, 30(2): 488-507.

de Onis, M., Wijnhoven, T., Onyango, A. (2004). Worldwide practices in child growth monitoring. *The Journal of Pediatrics*, 144 (4), p.461-465.

de Onis, M., Garza, C., Victora, C., Onyango, A., Frongillo, E., Martines, J. (2004).

The WHO Multicentre Growth Reference Study: Planning, study design, and methodology. *Food and Nutrition Bulletin*, vol. 25, no.1 (supplement).

de Onis, M., Onyango, A., Borghi, E., Siyam, A., Blössner, M., Lutter, C. (2012). Worldwide implementation of the WHO Child Growth Standards. *Public Health Nutrition*, 15(9), 1603-1610.

Deter, R., Harrist, R., Hadlock, F. and Poindexter, A. (1982). Longitudinal studies of fetal growth with the use of dynamic image ultrasonography. *American Journal of Obstetrics and Gynaecology*, 143:545-554.

Gallivan, S., Robson, S., Chang, T., Vaughn, J. and Spencer, J. (1993). An investigation of fetal growth using serial ultrasound data. *Ultrasound in Obstetrics and Gynaecology*, 3:109-114.

Gannoun, A., Girard, S., Guinot, C. and Saracco, J. (2002). Reference curves based on non-parametric quantile regression. *Statistics in Medicine*, 21:3119-3135.

Graber, E. (2021). *Physical Growth of Infants and Children*. [online] MSD Manuals. Available at: https://www.msdmanuals.com/professional/pediatrics/growth-and-development/physical-growth-of-infants-and-children?query=Physical%20Growth%20of%20Infants%20and%20Children. [Accessed 14 September 2021].

Hastie, T. and Tibshirani, R. (1986). Generalised Additive Models. *Statistical Science*, Vol. No. 3: 297-318.

He, X. (1997). Quantile Curves Without Crossing. *The American Statistician*, 51(2):186-192.

Koenker, R. and Basset, G. (1978). Regression Quantiles. *Econometrica*, 46(1): 33-50.

Muggeo, V., Sciandra, M., Tomasello, A. and Calvo, S. (2013). Estimating growth charts via non-parametric quantile regression: a practical framework with application in ecology. *Environmental and Ecological Statistics*, 20:519-531.

Natale, V. and Rajagopalan, A. (2014). Worldwide variation in human growth and the World Health Organisation growth standards: a systematic review. *BMJ Open*, 4:e003735

Persson, P. (1988). Fetal growth Curves. In Sharp, F., Fraser, R. and Milner, R. (eds.) *Fetal Growth*, 251-61.

Royston, P., Altman, D. (1994). Regression using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *Applied Statistics*, 43 No. 3: 429-467.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rigby, R. and Stasinopoulos, D. (2004). Smooth centile curves for skew and kurtotic data modelled using the Box-Cox Power Exponential Distribution. *Statistics in Medicine*, 23:3053-3076.

Rigby, R. and Stasinopoulos, D. (2005). Generalized additive models for location, scale and shape (with discussion). *Applied Statistics*, 54, part 3, pp 507-554.

Rigby, R. and Stasinopoulos, D. (2006). Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, 6: 209-229.

Rigby, R. and Stasinopoulos, D. (2010). *A Flexible approach using GAMLSS in R*.

Royston, P. (1995). Calculation of unconditional and conditional reference intervals for foetal size and growth from longitudinal measurements. *Statistics in Medicine*, 14: 1417-1436.

Royston, P., Altman, D. (1995). Design and analysis of longitudinal studies of fetal size. *Ultrasound in Obstetrics and Gynaecology*, 6: 307-312.

Schnabel, S. and Eilers, P. (2013). Simultaneous estimation of quantile curves using quantile sheets. *AStA Advances in Statistical Analysis*, 97(1): 77-87.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of Statistics*, 6:461-464.

Stasinopoulos, D. and Rigby, R. (2007). Generalized Additive Models for Location Scale and Shape (GAMLSS) in R. *Journal of Statistical Software*, vol. 23, 7.

Stasinopoulos, D., Rigby, R., Heller, G., Voudouris, V., De Bastiani, F. (2017). *Flexible Regression and Smoothing: Using GAMLSS in R*. USA: CRC Press.

Tarwa, C. and De Villiers, F. (2007). The use of the Road to Health Card in monitoring child health. *South African Family Practice*, 49:1, 15-15d.

Tilling, K., Macdonald-Wallis, C., Lawlor, D., Hughs, R. and Howe, L. (2014). Modelling Childhood Growth Using Fractional Polynomials and Linear Splines. *Annals of Nutrition and Metabolism*, 65: 129-138.

van Buuren, S. and Fredriks, M. (2001). Worm plot: a simple diagnostic device for modelling growth reference curves. *Statistics in Medicine*, 20:1259-1277.

van Buuren, S. (2007). Worm plot to diagnose fit in quantile regression. *Statistical Modelling*, 7(4): 363-376.

Verburg, B., Steegers, E., De Ridder, M., Snijders, R., Smith, E., Hofman, A., Moll, H., Jaddoe, V. and Witteman, J. (2008). New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound in Obstetrics and Gynaecology*, 31: 388-396.

Wade, A. and Ades, A. (1998). Incorporating correlations between measurements into the estimation of age-related reference ranges. *Statistics in Medicine*, 17: 1989-2002.

Wang, Y., Moreno, L., Caballero, B., Cole, T. (2006). Limitations of the current World Health Organisation growth references for children and adolescents. *Food and Nutrition Bulletin*, vol. 27, no. 4 (supplement).

Wen, X., Kleinman, K., Gillman, M., Rifas-Shiman, S. and Taveras, E. (2012). Childhood body mass index trajectories: modelling, characterising, pairwise correlations and socio-demographic predictors of trajectory characteristics. *BMC Medical Research Methodology*, 12:38

Wei, Y., Pere, A., Koenker, R. and He, X. (2006). Quantile regression methods for reference growth charts. *Statistics in Medicine*, 25(8): 1369-1382.

World Health Organization, WHO Multicentre Growth Reference Study Group. (2006). *WHO Child Growth Standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. Geneva: World Health Organization.

World Health Organization, WHO Multicentre Growth Reference Study Group. (2006). *WHO child growth standards: head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age: methods and development.* Geneva: World Health Organization.

World Health Organisation. (2006). *WHO Child Growth Standards: R igrowup package*. Geneva: World Health Organization.

World Health Organization. (2008). *Training Course on Child Growth Assessment*. Geneva: World Health Organisation.

Yang, Z., Duan, Y., Ma, G., Yang, X., Yin, S. (2015). Comparison of the China Growth Charts with the WHO growth standards in assessing malnutrition of children. *BMJ Open*, 5(2).

Zar, H., Barnett, W., Myer, L., Stein, D., Nicol, M. (2015) Investigating the early-life determinants of illness in Africa: The Drakenstein Child Health Study. *Thorax*, 70(6): 592-94