

Modelos Aditivos Generalizados para Posição, Escala e Forma (GAMLSS) na Modelagem de Curvas de Referência

Generalized Additive Models for Location, Scale and Shape (GAMLSS) for Modeling Reference

CLÁUDIO SÉRGIO MEDEIROS PAIVA¹
DJACYR MAGNA CABRAL FREIRE²
JOSÉ GUILHERME CECATTI³

RESUMO

Curvas de referência para valores de parâmetros antropométricos na população são amplamente utilizadas por profissionais de saúde. Técnicas estatísticas de construção de modelos matemáticos relacionadas com a idade desenvolveram-se bastante nos últimos 10 anos e tem sido uma área de especial interesse em vários centros de pesquisas no mundo. A evolução do poder de processamento dos computadores pessoais contribuiu de maneira significativa para esta mudança nos métodos de análise de regressão até então utilizados. A publicação do Estudo Multicêntrico de Referência do Crescimento da OMS (2006) criou um novo paradigma para construção de curvas de crescimento, e em consequência, implicou em um amplo questionamento a respeito das curvas anteriormente publicadas e atualmente em uso em várias áreas da saúde. O objetivo deste artigo foi expor de maneira concisa e sem aprofundamento dos aspectos matemáticos as bases metodológicas da modelagem estatística utilizando uma nova classe de modelos estatísticos chamada Modelos Aditivos Generalizados para Posição, Escala e Forma (GAMLSS) proposta por RIGBY e STASINOPOULOS (2005). Demonstrou-se a aplicação do método GAMLSS na elaboração de um modelo com dados empíricos, a utilização das ferramentas de diagnóstico e de seleção do modelo de melhor ajuste, a construção de uma curva de referência e o cálculo facilitado dos percentis e z-escores através de funções específicas. Conclui-se que o método GAMLSS é uma importante ferramenta de modelagem estatística na área da saúde e que em futuras pesquisas provavelmente será o estado da arte na construção de novas curvas de referência ou na validação de curvas já construídas.

DESCRIPTORIOS

Métodos Epidemiológicos. Modelos Estatísticos. Análise de Regressão

ABSTRACT

Reference curves for values of anthropometric parameters of a population are widely used by healthcare professionals. Statistical techniques for the construction of age-related mathematical models have developed a lot in the last 10 years and this has been an area of particular interest in several research centers around the world. The evolution of the processing power of personal computers has contributed significantly to this change in the methods of regression analysis used until then. The WHO publication of the Multicenter Growth Reference Study (MGRS) in 2006 created a new paradigm for the construction of growth curves, which, consequently, resulted in a broad discussion on the curves previously published and currently in use in several healthcare areas. The aim of this exhibit was to expose in a concise way — without further consideration of the mathematical aspects — the methodological bases of statistical modeling using a new class of statistical models called Generalized Additive Models for Position, Scale and Shape (GAMLSS) proposed by RIGBY and STASINOPOULOS (2005). The application of the GAMLSS method to design a model with empirical data was demonstrated, as well as the use of diagnostic tools and the selection of the model of best fit, the construction of a curve of reference and easier calculation of percentiles and z-scores through specific functions. It is concluded that the GAMLSS method is an important tool for statistical modeling in healthcare and that in future research it will probably be the state of the art in the construction of new reference curves or in the validation of curves already built.

DESCRIPTORS

Epidemiological Methods. Statistical Models. Regression Analysis.

1 Prof. Adjunto IV do Departamento de Obstetrícia e Ginecologia do Centro de Ciências Médicas. Universidade Federal da Paraíba. João Pessoa- Paraíba-Brasil.

2 Prof^º. da Escola Técnica de Saúde do Centro de Ciências da Saúde, Universidade Federal da Paraíba;

3 Professor do Departamento de Tocoginecologia da Faculdade de Ciências Médicas da UNICAMP. Campinas-São Paulo.

Curvas de referência para valores de parâmetros antropométricos na população são amplamente utilizados por profissionais da área de saúde para comparar uma medição de um único indivíduo com o conjunto de valores observados em um controle populacional. Intervalos de referência ajustados para alguma covariável podem ser utilizados para avaliar indivíduos em um único ponto no tempo ou para monitorar mudanças no mesmo indivíduo ao longo do tempo. Nas curvas de crescimento, a covariável mais comumente utilizada é a idade do indivíduo (COLE e GREEN, 1992).

Em aplicações na área de saúde muitas vezes é necessário ajustar centis de referência para a idade do indivíduo considerado saudável. As curvas de crescimento típicas destas aplicações constituem gráficos que representam alguns quantis selecionados da distribuição de alguma característica física em uma população de referência em função da idade dos indivíduos (DIBLEY *et al.*, 1987). Os gráficos de referência ou curvas de normalidade são ferramentas clínicas de aplicações muito variadas. Como por exemplo, são utilizadas para monitorar o crescimento individual de uma criança, para estimar a idade gestacional a partir da biometria fetal pela ultrassonografia, para classificar e selecionar indivíduos para investigação clínica ou intervenção terapêutica quando alguma característica destes se desvia do padrão observado numa população saudável, ou como indicador de saúde pública para comparar o estado nutricional ao longo do tempo ou mesmo entre diferentes populações.

Técnicas estatísticas de construção de modelos matemáticos relacionadas com a idade desenvolveram-se bastante nos últimos 10 anos e tem sido uma área de especial interesse em vários centros de pesquisas no mundo. A evolução do poder de processamento dos computadores pessoais contribuiu de maneira significativa para esta mudança nos métodos de análise de regressão até então utilizados.

Os métodos para a construção de curvas centílicas de base populacional covariável-ajustados foram recentemente revistos pela Organização Mundial da Saúde (OMS) e estão agora bem estabelecidos (BORGHI *et al.*, 2006; WHO, 2006). Em abril de 2006, a OMS liberou o “Padrão de crescimento infantil para crianças de 0 a 5 anos” que foi gerado pelo “Estudo Multicêntrico de Referência do Crescimento da OMS” (EMRC). A publicação deste estudo criou um novo paradigma para construção de curvas de crescimento e, como consequência, implicou em um amplo questionamento a respeito das curvas anteriormente publicadas e atualmente em uso em várias áreas da saúde.

Reference curves for values of anthropometric parameters of a population are widely used by healthcare professionals to compare the measures of one single individual with the set of values observed in one control population. Reference intervals fit to some co-variable can be used to evaluate individuals in one single moment or to monitor changes in the same individual along time. When it comes to growth curves, the most commonly used co-variable is the individual's age (COLE and GREEN, 1992).

For applications in healthcare, it is usually necessary to adjust reference centiles to the age of the individual considered healthy. Typical growth curves for these applications constitute charts that represent some selected quantiles of the distribution of some physical characteristic in one reference population in function of the individuals' age (DIBLEY *et al.*, 1987). The reference charts or normality curves are clinical tools of varied application. As an example, they are used to monitor the individual growth of a child, to estimate gestational age from fetal biometry by ultrasound scan, to classify and select individuals for one clinical investigation or therapeutic intervention when any of their characteristic deviates from the standard observed in a healthy population, or as a public health indicator to compare nutritional state throughout time or even among different populations.

Statistical techniques for the construction of age-related mathematical models have developed a lot in the last 10 years and this has been an area of particular interest in several research centers around the world. The evolution of the processing power of personal computers has contributed significantly to this change in the regression analysis methods used until then.

The methods for the construction of population-based centile curves were recently reviewed by the World Health Organization (WHO) and are now well established (BORGHI *et al.*, 2006; WHO, 2006). In April 2006, WHO released the “Child growth standards for children between 0 and 5 years old”, which was generated by the “WHO Multicenter Growth Reference Study (MGRS). The publication of this study set a new paradigm for growth curves construction and, consequently, resulted in a broad discussion on the curves previously published and currently in use in several healthcare areas.

As part of a vast consultative process to select the best statistical methods for the construction of new

Como parte de um amplo processo consultivo para selecionar os melhores métodos estatísticos para a construção de novas curvas de referência, em janeiro de 2003 a OMS convocou um grupo de peritos em modelagem estatística que discutiu as questões metodológicas enfrentadas no processo de construção das novas curvas de crescimento e avaliou os pontos fortes e fracos dos vários métodos disponíveis. O grupo analisou 30 métodos estatísticos diferentes e recomendou a utilização do método BoxCox Power Exponencial (BCPE) com alisamento das curvas utilizando splines cúbicas (BORGHI *et al.*, 2006). Este método faz parte de uma ampla classe de modelos estatísticos chamada GAMLSS (Modelos Aditivos Generalizados para Posição, Escala e Forma) proposta por RIGBY e STASINOPOULOS (2005). Esta classe é uma evolução da classe GLM (Modelos Lineares Generalizados) e consiste basicamente em uma teoria de regressão para variáveis dependentes da família exponencial de distribuições. O GAMLSS é um pacote de expansão do software estatístico R, de livre acesso e disponibilizado gratuitamente para a comunidade científica no endereço eletrônico <http://www.gamlss.com/>.

Este artigo pretende expor para os profissionais da área de saúde, tanto assistenciais quanto potenciais pesquisadores, de maneira concisa e sem aprofundamento dos aspectos matemáticos, as recomendações emanadas desse grupo consultivo da OMS (WHO, 2006).

MÉTODO

Aspectos gerais da modelagem estatística

Segundo MASSAD (2004) um “*Modelo*” pode ser definido como uma representação conveniente de alguma coisa importante, e quando esta representação consiste de componentes quantitativos é chamado de *modelo matemático*. O processo de modelagem consiste em uma série atividades complexas associadas ao desenho de modelos representando o mundo real e sua solução. Um *modelo matemático* é uma representação ou interpretação simplificada da realidade descrita em linguagem matemática, mediante simplificações do sistema (BRIASSOULIS, 2008).

Para RIGBY e STASINOPOULOS (2008), *modelagem estatística* é a arte de construir modelos estatísticos parcimoniosos para uma melhor compreensão dos fenômenos de interesse. Deve-se considerar que um bom modelo deverá ter pelo menos estas quatro qualidades fundamentais: parcimônia, generalidade,

reference curves, in January 2003, WHO called a group of experts in statistical modeling to discuss the methodological issues faced in the process of building new growth curves and to evaluate the weak and strong points of various methods already available. The group analyzed 30 different statistical methods and recommended the use of the BoxCox Power Exponential method (BCPE) to smooth the curves using cubic splines (BORGHI *et al.*, 2006). This method made part of a broad class of statistical models called GAMLSS (Generalized Additive Models for Location, Scale and Shape) proposed by RIGBY and STASINOPOULOS (2005). This class is an evolution of the GLM class (Generalized Linear Models) and consists basically of a regression theory for variables that depend on the exponential family of distributions. GAMLSS is an expansion package of the R statistical software, of free access and made available for the scientific community at <http://www.gamlss.com/>.

This paper aims at exposing to healthcare professionals, both physicians and researchers, in a concise way and without further discussion on mathematical aspects, the recommendations which resulted from this WHO consultative group (WHO, 2006).

METHOD

General aspects of statistical modeling

According to MASSAD (2004) one “*Model*” can be defined as a convenient representation of something important, and when this representation consists of quantitative components it is called *mathematical model*. The modeling process consists of a series of complex activities associated to the design of models that represent the real world and its solution. One mathematical *model* is a representation or simplified interpretation of reality described in mathematical language, based on system’s simplifications (BRIASSOULIS, 2008).

To RIGBY and STASINOPOULOS (2008), *statistical modeling* is the art of constructing parsimonious statistical models for a better comprehension of phenomena of interest. It must be understood that a good model shall have at least these

consiliência (em unidade com o conhecimento consolidado) e capacidade preditiva (MASSAD, 2004).

Em sintonia com o princípio da parcimônia de *Ockham*, um modelo deve ser uma simplificação da realidade, pois representam uma versão “suavizada” da realidade, mais fácil de compreender e de interpretar. Modelos não têm de ser verdadeiros (na verdade eles geralmente não são), mas eles têm de ser úteis para a finalidade para a qual foram construídos. Um bom modelo pode ser usado para uma finalidade explicativa, para testar hipóteses e para fazer previsão. Dentro da classificação dos modelos com vistas ao propósito de sua aplicação, classificam-se os modelos de simulação de acordo com os seguintes propósitos: *modelos voltados à previsão; modelos voltados à investigação e modelos voltados à comparação*. As curvas de crescimento focalizadas nesse artigo são exemplos de modelos voltados à comparação.

Métodos de modelagem estatística para a construção de curvas de referência

Para elaborar um modelo estatístico e construir uma curva de referência há várias questões preliminares que precisam ser respondidas. Como é que a média de y depende de x ? Como é que a variância de y depende de sua média? Qual é a distribuição de y dado x ? Será que os parâmetros de escala e forma da distribuição de y dependem de x ? (STASINOPOULOS E RIGBY, 2007).

O passo seguinte é verificar, após uma análise da distribuição dos dados, se ao longo do intervalo da variável explanatória (no caso de curvas de crescimento, a idade) a forma da dispersão dos dados admite os pressupostos da distribuição normal ou de outra família de distribuição conhecida. Distribuições geralmente são caracterizadas por medidas resumo relacionadas com três momentos: a média, o desvio-padrão (ou coeficiente de variação) e a assimetria. Várias medidas antropométricas (como peso, estatura, índice de massa corporal, perímetro cefálico, espessura da prega subescapular, índices dopplervelocimétricos, etc.) em muitas publicações científicas costumam ser assumidas como tendo distribuição normal. Porém, outros estudos (AKANTZILIOTOU *et al.*, 2002; RIGBY e STASINOPOULOS, 2004; WHO, 2006) demonstraram que muitas dessas variáveis apresentam distribuições com assimetria e, algumas vezes, curtose. O efeito do quarto momento da distribuição (curtose) é significativo na estimativa dos percentis extremos (VAN BUUREN e FREDRIKS, 2001; RIGBY e STASINOPOULOS, 2004), como por exemplo, os 3^o, 5^o, 95^o, 97^o percentis, precisamente aqueles mais utilizados

four fundamental qualities: parsimony, generality, consilience (in conformity with the consolidated knowledge) and predictive capacity (MASSAD, 2004).

In sync with *Ockham's* principle of parsimony, one model must be one simplification of reality, for it represents one “softened” version of reality, easier to comprehend and to interpret. Models do not have to be real (actually, they generally are not), but they have to be useful for the finality to which they were built. One good model can be used for one explanatory end, to test hypothesis and to make a prediction. In the model classification, as to its application, simulation models are classified according to the following purposes: *prediction models; investigation models and comparison models*. Growth curves studied in this article are examples of comparison models.

Statistical modeling methods for building reference curves

To create a statistical model and build one reference curve there are various preliminary questions that need to be answered. How come does y 's mean depend on x ? How come does y 's variance depend on its mean? What is the distribution of y with a given x value? Do y 's scale and form distribution parameters depend on x ? (STASINOPOULOS and RIGBY, 2007).

The next step is to verify, after analyzing the data distribution, if — throughout the explanatory variable interval (age, in the case of growth curves) — the form of data dispersion admits the presupposition of normal distribution or of another distribution family already known. Distributions are generally characterized by summary measures related to three moments: mean, standard deviation (or variation co-efficient) and asymmetry. Several anthropometric measures (such as weight, height, body mass index, skull perimeter, subscapular skin fold thickness, Doppler velocimetry indexes, etc.) in several scientific publications are generally assumed as having normal distribution. However, other studies (AKANTZILIOTOU *et al.*, 2002; RIGBY and STASINOPOULOS, 2004; WHO, 2006) demonstrate that a lot of these variables present asymmetric distributions and, sometimes, kurtosis. The effect of the fourth moment of distribution (kurtosis) is meaningful in the estimate of the extreme percentiles (VAN BUUREN and FREDRIKS, 2001; RIGBY and STASINOPOULOS, 2004), as, for example, the 3rd, 5th, 95th, 97th percentiles; precisely those which are most used

como limites de normalidade em muitas curvas de crescimento utilizadas na área de saúde. As atuais recomendações da OMS para a construção de novas curvas de crescimento exigem que se verifique se há adequada concordância entre a distribuição empírica dos dados e os pressupostos da distribuição selecionada pelo método (WHO, 2006; ONIS *et al.*, 2007).

Outra questão prioritária na tarefa de modelar curvas de crescimento é saber qual foi o desenho da coleta dos dados. Os métodos aplicados comumente assumem o desenho transversal, isto é, que as medidas utilizadas para a construção do modelo são independentes. Neste caso, considera-se a utilização de modelos lineares ou não-lineares de efeitos fixos. Se o desenho for longitudinal ou mesmo combinado (como foi o caso da curva da OMS (WHO, 2006) que teve um componente longitudinal e outro transversal), o conjunto de dados contém medidas seriadas de um mesmo elemento amostral. Como estes dados estão autocorrelacionados, o pressuposto de independência dos dados não é satisfeita. Quando se está lidando com medidas repetidas, são particularmente apropriados os modelos de efeitos aleatórios ou de efeitos mistos. Estes modelos de efeitos mistos podem ser utilizados para modelar curvas de crescimento, uma vez que permitem a partição da variância entre dois componentes: a contribuição da variabilidade entre os sujeitos e a variabilidade intra-sujeitos.

BORGHI *et al.* (2006) encontraram pelo menos 30 métodos que poderiam potencialmente ser utilizados na construção da curva de crescimento do EMRC (WHO, 2006). Classificaram os métodos com base em dois critérios: a) tratamento da variável idade (se agrupavam as idades ou tratavam a idade como variável contínua) e b) utilização ou não de pressupostos da distribuição dos dados. A seguir, subdividiram os métodos em quatro categorias: a) métodos de agrupamento e alisamento, sem pressupostos de distribuição; b) métodos de agrupamento e alisamento, com pressupostos de distribuição; c) métodos que manipulam a idade continuamente, sem pressupostos de distribuição; d) métodos que manipulam a idade continuamente, com pressupostos de distribuição. No quadro 1 estão listados alguns métodos estatísticos selecionados do estudo de BORGHI *et al.* (2006).

Para a construção do “Padrão de crescimento infantil para crianças de 0 a 5 anos” o EMRC da OMS selecionou o modelo Box-Cox-power-exponencial (BCPE) de RIGBY e STASINOPOULOS (2004), que pertence à família dos modelos aditivos generalizados para posição, escala e forma (GAMLSS) e fornece

as normality limits in several growth curves used in healthcare. WHO’s current recommendations for the construction of new growth curves demand that it be verified if there is adequate agreement between the empirical distribution of data and the presuppositions of distribution selected by the method (WHO, 2006; ONIS *et al.*, 2007).

Another priority question in the task of modeling growth curves is to know how the data collection was done. The applied methods commonly take the transversal design, that is, the measures used for the construction of the model are independent. In this case, we consider the use of linear models or non-linear of fixed effects. If the design is longitudinal or even combined (as was the case of the WHO curve (WHO, 2006), which had one longitudinal and another transversal component), the set of data contains serial measures of a single sample element. Since these data are self-correlated, the data independence presupposition is not satisfied. When one is dealing with repeated measures, random effect or mixed effect models are particularly appropriate. These mixed effect models can be used to model growth curves, since they allow the participation of variance between two components: the contribution of variability between the individuals and the intra-individual variability.

BORGHI *et al.* (2006) found at least 30 methods that could potentially be used in the construction of the MGRS growth curve (WHO, 2006). They classified the methods based on two criteria: treatment of the age variable (ages were grouped or treated as a continuous variable) and b) use or not of data distribution presupposition. Next, they subdivided the methods in four categories: a) grouping and smoothing methods, without distribution presuppositions; b) grouping and smoothing methods, with distribution presuppositions; c) methods that manipulate age continuously, without distribution presuppositions; d) methods that manipulate age continuously, with distribution presuppositions. Chart 1 lists some statistical methods selected in the study of BORGHI *et al.* (2006).

For the construction of “Child growth standards for children between 0 and 5 years old” WHO’s MGRS selected the Box-Cox-power-exponential model (BCPE) by RIGBY and STASINOPOULOS (2004), which belongs to the family of Generalized Additive Models for Location, Scale and Shape (GAMLSS) and provides highly flexible models for various types of distributions

Quadro 1 - Classificação dos métodos de construção de curvas de crescimento segundo o tratamento da variável idade, método de ajuste da curva e pressupostos da distribuição(*).

Chart 1 – Classification of the growth curves construction methods according to the age variable treatment, curve fitting method and distribution presuppositions (*)

Método Method	Tratamento da variável idade Age variable Treatment	Estimação dos centis Centile estimate	Método de ajuste da curva Curve fitting method	Pressupostos da distribuição Distribution presuppositions
Splines com <i>knots</i> fixos Fixed knot splines	AA	S	Splines com <i>knots</i> fixos	Nenhum
Ajuste visual Visual fitting	AA	S	Ajuste visual	Nenhum
Regressão de núcleo Nucleus regression	AA	S	Ajuste visual	Regressão de núcleo
Ajuste visual Visual fitting	AA	J	Ajuste visual	Normal
Box-Cox normalização Box-cox normalization	AA	J	Ajuste visual	Box-Cox normal
LMS (versão 1) LMS (version 1)	AA	J	<i>Splines</i> cúbicas	Box-Cox normal
Estabilização de variâncias Variance Stabilization	AA	J	Polinômios	Normal
Ajuste polinomial Polynomial fitting	AA	J	Polinômios	Normal
Regressão quantílica Quantile Regression	MC	J	<i>Splines</i> naturais	Nenhum
Curvas quantílicas não-Gaussianas Non Gaussian Quantile Curves	MC	J	Função de 4 parâmetros	Nenhum
Modelagem multinível Multilevel modeling	MC	J	Estimação de modelos lineares e não-lineares	
LMS (versão 2) LMS (version 2)	MC	J	<i>Splines</i> cúbicas	Box-Cox normal
Aditividade e estabilização da variância (AVAS) Additivity and Variance Stabilization (AVAS)	MC	J	Regressão não-paramétrica	Normal
GAMLSS GAMLSS	MC	J	Funções paramétricas ou não- paramétricas	Box-Cox Power exponencial
LMST	MC	J	<i>Splines</i> cúbicas ou polinômios fracionais	Box-Cox t
LMSP LMSP	MC	J	<i>Splines</i> cúbicas ou polinômios fracionais	Box-Cox Power exponencial

(*) adaptado de Borghi *et al.*, 2006; AA – agrupamento e alisamento; MC – manipulada continuamente; S – separadamente; j - juntos

(*) adapted from Borghi *et al.*, 2006; GS – Grouping and Smoothing; CM – Continuous Manipulation; S – Separate; T – Together

modelos altamente flexíveis para vários tipos de distribuições com diferentes características para a posição, escala, assimetria e curtose da variável resposta. O BCPE é uma família de distribuição que incorpora quatro parâmetros, sendo que cada um pode ser modelado como função paramétrica (por exemplo, polinômios fracionais) ou função não-paramétrica (por exemplo, *splines* cúbicas) da variável explanatória.

with different characteristics for location, scale, asymmetry and kurtosis of varied response. BCPE is a distribution family that incorporates four parameters; each one can be modeled as a parametric function (for example, fractional polynomials) or non-parametric function (for example, cubic *splines*) of the explanatory variable.

Os modelos aditivos generalizados para posição, escala e forma foram apresentados inicialmente por RIGBY e STASINOPOULOS (2001) como uma abordagem para superar algumas limitações dos Modelos Lineares Generalizados (MLG) e dos Modelos Aditivos Generalizados (MAG). Sua proposta é ser um método para ajuste de qualquer tipo de modelo de regressão, pois constituem uma evolução dos MLG e MAG.

Por sua vez, os MAG são uma extensão do modelo linear generalizado em que o termo aditivo é substituído por uma função não paramétrica (isto é, cuja forma não é especificada) da variável explicativa, estimada através de curvas de alisamento a partir do conjunto de dados (HASTIE e TIBSHIRANI, 1987). A finalidade dos MAG é maximizar a qualidade de predição de uma variável dependente Y calculando funções inespecíficas (não-paramétricas) das variáveis preditoras que estão “ligadas” com a variável dependente.

O GAMLSS é uma classe de modelos bastante flexível porque permite além de escolher entre uma grande variedade de distribuições para a variável dependente, escolher várias funções de ligação para os efeitos das variáveis preditoras sobre a variável dependente. Essa função estimada, também chamada de curva alisada ou *smoother*, pode assumir diversas formas, uma vez que não apresenta a estrutura rígida de uma função paramétrica. São exemplos de alisadores as *splines* cúbicas, as loess (locally weighted running line smoother) e os polinômios fracionais.

Uma *spline* é uma curva definida matematicamente por dois ou mais *pontos de controle*. Os pontos de controle que ficam na curva são chamados de *nós* (*knots*). A *spline* polinomial consiste em uma função de polinômios fracionados e emendados em nós. O termo *spline* vem de um dispositivo usado por antigos construtores de navios para desenhar formas mais suaves. Originalmente, uma *spline* era uma ferramenta utilizada por artesãos que consistia em uma haste fina e longa de madeira ou metal posicionada em vários pontos de fixação selecionados para dar a forma de uma curva lisa que passa através de pontos (nós).

Para escolher um modelo estatístico que se ajuste bem aos dados, um pesquisador precisa tomar algumas decisões análogas àquelas dos artesãos na construção de suas *splines* mecânicas. Os segmentos curvos da haste de madeira eram ajustados aos pontos de fixação com molas elásticas ou pesos. Se as molas fossem suficientemente fracas, a haste retornaria a ser reta (modelo de regressão linear) ou se fossem suficientemente fortes resultaria num aspecto serrilhado (interpolação linear). O alisamento mecânico da *spline*

Generalized Additive Models for Location, Scale and Shape were first presented by RIGBY and STASINOPOULOS (2001) as an approach to overcome some limitations of the Generalized Linear Models (GLM) and the Generalized Additive Models (GAM). Their proposition is to be a method to adjust any type of regression model, for it constitutes one evolution to GAM and GLM.

On its turn, GAM is one extension to the generalized linear model in which the additive term is substituted by one non-parametric function (that is, whose form is not specified) of the explanatory variable, estimated through smoothing curves from a data set (HASTIE and TIBSHIRANI, 1987). GAM'S goal is to maximize the quality of prediction of one dependent variable Y , by calculating unspecified (non-parametric) functions of the predictive variables that are “connected” with the dependent variable.

GAMLSS is one class of fairly flexible models because it allows — besides choosing from a wide range of distributions for the dependent variable — choosing various connection functions for the effects of predictive variables over the dependent variable. This estimated function, called smoother curve, can take up different forms, since it does not present the rigid structure of a parametric function. Cubic *splines*, loess (locally-weighted running-line smoother) and fractional polynomials are examples of such.

One *spline* is a curve mathematically defined by two or more *control points*. The control points that stay on the curve are called knots. The polynomial *spline* consists of one function of fractioned polynomials mended by knots. The term *spline* comes from a device used by ancient ship makers to design smoother forms. Originally, one *spline* was a tool used by craftsmen and it was made of a thin long wooden or metal stick placed on several preselected fixation points in order to give a smooth curvy shape which passes through points (knots).

In order to choose one statistical model that fits well to the data, one researcher needs to make some similar decisions to those the craftsmen did in the construction of their mechanical *splines*. The curvy segments of the wooden stick were adjusted to the fixation points with elastic springs or weights. If the springs were sufficiently weak, the stick would go back to the straight position (linear regression model); if they were sufficiently strong, a jagged aspect would result

fornece um contrabalanço entre fidelidade aos dados e alisamento, que depende da força das molas. Do ponto de vista estatístico, a grau de alisamento na modelagem de uma *spline* pode ser formalizado em um critério matemático: critério = resíduos do modelo + (contrabalanço x penalidade do alisamento), a ser explanado adiante.

No GAMLSS, a variável resposta $y \sim D(\mu, \sigma, \nu, \tau)$, onde D pode ser qualquer distribuição (incluindo distribuições contínuas altamente assimétricas e curtóticas e distribuições discretas) e as letras gregas μ (mu), σ (sigma), ν (nu) e τ (tau) representam os parâmetros da distribuição. Existem cerca de 40 tipos diferentes de famílias de distribuições disponíveis no GAMLSS, com um, dois, três ou quatro parâmetros, tanto distribuições contínuas (normal, lognormal, exponencial, gamma, beta, BoxCox power exponencial, BoxCox t, BoxCox Cole Green, Gumbell, Johnson's SU, Weibull, etc.), como distribuições discretas (Poisson, Sichel, binomial, beta binomial, etc.) ou mistas (Poisson zero inflacionada, Gaussiana inversa zero ajustada, beta inflacionada de zero e um, etc.). Os dois primeiros parâmetros da distribuição da população M e σ geralmente são caracterizados como parâmetros de localização e escala, enquanto os dois restantes parâmetros ν e τ , se existirem, são caracterizados como parâmetros de forma, por exemplo, assimetria e curtose. Todos os parâmetros da distribuição podem ser modelados como funções paramétricas lineares/não-lineares e/ou funções não-paramétricas suavizadas de variáveis explicativas (isto é, *splines* cúbicas, *splines* penalizadas, loess) e/ou efeitos randômicos (STASINOPOULOS E RIGBY, 2007).

Avaliação e Seleção dos modelos

Comparar diferentes modelos é a questão mais importante em modelagem estatística. Com o avanço da capacidade de processamento dos atuais computadores pessoais, desenvolvimento de novos softwares estatísticos e acessibilidade desses softwares à comunidade científica, múltiplos modelos mais complexos podem ser simultaneamente ajustados aos dados em análise. Isto implica que a modelagem estatística é um processo iterativo, pois os modelos devem ser construídos gradualmente e comparados entre si para a escolha do mais parcimonioso entre os de melhores ajustes.

Além disso, a variância dos dados em torno da média (ou mesmo da mediana) nem sempre pode ser admitida como uniforme ao longo do intervalo da variável explanatória (homocedasticidade). O principal

(linear interpolation). The mechanical smoothing of the *spline* supplies a counterbalance between fidelity to the data and smoothing, which depends on the force of the springs. From the statistical point of view, the degree of smoothing in the modeling of a *spline* can be formalized in one mathematical criterion: criterion = model residue + (counterbalance X smooth penalty), to be explained ahead.

In GAMLSS, the response variable $y \sim D(\hat{\mu}, \hat{\sigma}, \hat{\nu}, \hat{\tau})$, where D can be any distribution (including highly asymmetrical continual distributions, kurtosis, and discrete distributions) and the Greek letters $\hat{\mu}$ (mu), $\hat{\sigma}$ (sigma), $\hat{\nu}$ (nu) and $\hat{\tau}$ (tau) represent the distribution parameters. There are about 40 different types of distribution families available at GAMLSS, with one, two, three or four parameters; both continuous distributions (normal, lognormal, exponential, gamma, beta, BoxCox power exponential, BoxCox t, BoxCox Cole Green, Gumbell, Johnson's SU, Weibull, etc.) and discrete distributions (Poisson, Sichel, binomial, beta binomial, etc.), or mixed (zero-inflated Poisson, zero-adjusted inverse Gaussian, zero- and one-inflated beta, etc.). The two first distribution parameters of the M population and $\hat{\sigma}$ generally are characterized as location and scale parameters, whereas the other two parameters $\hat{\nu}$ and $\hat{\tau}$, if present, are characterized as form parameters, for example, asymmetry and kurtosis. All the distribution parameters can be modeled as linear/non-linear parametric functions and/or smooth non-parametric functions of explanatory variables (that is, cubic *splines*, penalized *splines*, loess) and/or random effects (STASINOPOULOS and RIGBY, 2007).

Evaluation and Selection of Models

Comparing different models is the most important question in statistical modeling. With the advances of the processing capacity of current personal computers, development of new statistical software and the access to these software by the scientific community, more complex multiple models can be simultaneously adjusted to the data under analysis. This implies that the statistical modeling is an iterative process, for the models must be constructed gradually and compared to one another in order to choose the most parsimonious among those that fit best.

Besides this, data variance around the mean not always can be admitted as uniform along the interval of the explanatory variable (homoscedasticity). The main challenge in the construction of centile curves is to

desafio na construção de curvas centílicas é suavizar a distribuição em duas direções simultaneamente, ou seja, tanto entre as idades quanto dentro das idades e definir a quantidade de alisamento que promove um razoável contrabalanceamento entre parcimônia do modelo e fidelidade aos dados (VAN BUUREN E FREDRICKS, 2001).

Atualmente, a modelagem por análise de regressão linear e ajuste pelo método dos mínimos quadrados ainda é a técnica mais utilizada nas publicações científicas que envolvem modelagem estatística. Isto tem ocorrido não só pelo fato de menor exigência computacional no cálculo dos parâmetros do modelo, mas principalmente pela parcimônia e pela maior facilidade de entendimento do modelo gerado. No entanto, modelos aditivos generalizados são muito mais flexíveis e podem proporcionar um excelente ajuste na presença de relacionamentos não-lineares. Por outro lado, devido a essa flexibilidade, o pesquisador deve ser cauteloso na escolha do modelo para não sobreajustar os dados, ou seja, aplicar um modelo excessivamente complexo (com vários graus de liberdade) de modo a produzir um bom ajuste que provavelmente não será replicado em estudos de validação posteriores.

Ferramentas de diagnóstico

Existem diversas ferramentas estatísticas para realizar um diagnóstico da adequação do ajuste de um modelo e para selecionar o modelo mais adequado. VAN BUUREN e FREDRIKS (2001) destacaram as seguintes ferramentas: inspeção visual da forma das curvas de referência; percentis plotados contra os valores dos dados individuais; percentis empíricos e percentis ajustados plotados juntos; verificação da proporção dos dados com valores acima e abaixo dos percentis ajustados; testes estatísticos para verificar a normalidade dos resíduos; gráfico quantil-quantil dos z-escores (“*qq-plot*” e “*detrended qq-plot*”); o gráfico “*worm plot*” de VAN BUUREN e FREDRIKS (2001); desvio global (*deviance*), critério de informação de Akaike (AIC), critério bayseano de Schwartz (SBC) e o teste Q de Royston e Wright (2000). Não é do escopo deste artigo detalhar a fundamentação estatística, vantagens e limitações de cada ferramenta. Não há uma regra para a escolha das ferramentas nem em que sequência devem ser utilizadas, porém as cinco últimas ferramentas são as mais importantes disponíveis no pacote estatístico GAMLSS e por isto as comentaremos.

A análise dos resíduos é facilmente realizada no GAMLSS pela função *plot()*. Esta função produz quatro gráficos de resíduos quantílicos normalizados (resíduos

smooth the distribution in two ways simultaneously, that is, both among ages and within ages, and define the quantity of smoothing that promotes one reasonable counterbalance between model parsimony and data fidelity (VAN BUUREN E FREDRICKS, 2001).

Currently, modeling by linear regression analysis and adjusting by the least squares method still is the most used technique in scientific publications which involve statistical modeling. This has happened not only because of less computational demand in the calculus of the model parameters, but mainly because of parsimony and by the greater ease of understanding the generated model. However, generalized additive models are much more flexible and can provide one excellent fit in the presence of non-linear relationships. On the other hand, due to this flexibility, the researcher must be careful when choosing the model so as not to overfit the data, that is, to apply one excessively complex model (with several degrees of freedom) so as to produce good fitting that probably will not be replicated in further validation studies.

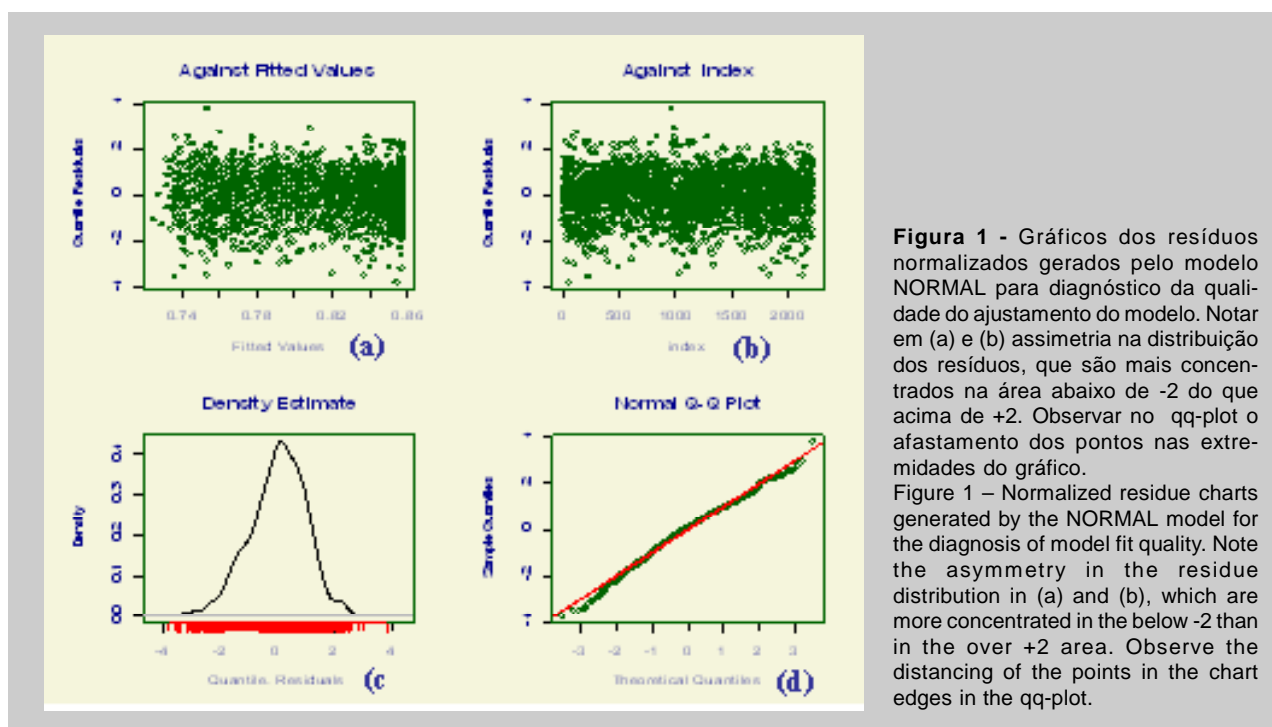
Diagnostic Tools

There are several statistical tools to diagnose the adequacy of model adjustment and to select the most adequate model. VAN BUUREN and FREDRIKS (2001) point out the following tools: visual inspection of the reference curve form; percentiles plotted against the values of the individual data; empirical percentiles and fit percentiles plotted together; data ratio verification with values over and below the fit percentiles; stylistic tests to verify the normality of residues; z-score quantile-quantile chart (“*qq-plot*” and “*detrended qq-plot*”); “*worm plot*” chart by VAN BUUREN and FREDRIKS (2001); global deviance, Akaike information criterion (AIC), Schwartz Bayesian criterion (SBC) and Q-test by Royston and Wright (2000). This article does not aim at pinpointing the statistical fundamentals, advantages and limitations of each tool. There is not a rule for the choice of tools or in which sequence they must be used; however, the five last tools are the most important available in the GAMLSS statistical package and, for this, they will be commented upon.

The residue analysis is easily done in the GAMLSS by the *plot()* function. This function

contra valores ajustados; resíduos contra variável explanatória; função densidade dos resíduos; *qq-plot* dos resíduos e um sumário de medidas resumo da distribuição dos resíduos, em que informa a média, variância, coeficiente de assimetria e coeficiente de curtose, permitindo verificar o pressuposto da normalidade dos resíduos (Figura 1). Assim, se um modelo se ajustar bem aos dados, os seus resíduos verdadeiros deverão apresentar distribuição aproximadamente normal padrão, ou seja, média próximo de zero, variância próximo de 1, coeficiente de assimetria próximo de 0 e coeficiente de curtose próximo de 3, mesmo quando a distribuição do modelo não é normal (DUNN e SMITH, 1996). A normalidade dos resíduos também poderá ser verificada pela interpretação dos gráficos de dispersão dos resíduos contra a variável explanatória ou contra valores ajustados e transformados, de tal forma que se tenha uma variância constante para a distribuição utilizada. O padrão nulo desses gráficos é uma distribuição dos resíduos em torno de zero com amplitude constante. A análise gráfica residual permite verificar afastamentos da normalidade dos resíduos.

produces four charts of normalized quantile residue (residues against fit values; residues against explanatory variable; residue density function, residue *qq-plot* and one measure summary of residue distribution, in which mean, variance, asymmetry coefficient and kurtosis coefficient are informed, thus, allowing verifying the presupposition of residue normality (Figure 1). Thus, if one model fits well to data, its true residues shall present a nearly normal standard distribution, that is, mean near zero, variance near 1, asymmetry coefficient near 0 and kurtosis coefficient near 3, even when the distribution of the model is not normal (DUNN and SMITH, 1996). Residue normality can also be verified by the interpretation of residue dispersion charts against the explanatory variable or against adjusted and transformed values, in such a way that we have a constant variance for the used distribution. These charts' null standard is a residue distribution around zero with constant amplitude. The residue chart analysis allows verifying distancing of the residue normality.



Se os resíduos são normalmente distribuídos, então alguns pontos irão cair acima da linha horizontal, outros irão cair abaixo dela, e o padrão dos escores acima e abaixo da linha horizontal será aleatório. A interpretação visual é subjetiva e depende da experiência

If the residues are normally distributed, then, some points will be seen over the horizontal line, others will be seen below it, and the score standard over and below the horizontal line will be random. The visual interpretation is subjective and it depends on the

do pesquisador, porém uma regra básica na interpretação dos gráficos de resíduos é que o aparecimento de uma nuvem de pontos com uma forma clara em vez de um padrão aleatório é um aviso de que os resíduos não são normalmente distribuídos. A etapa seguinte é buscar apoio nos estatísticos de normalidade.

A normalidade dos resíduos também poderá ser verificada por testes estatísticos específicos, como por exemplo, o teste de Kolmogorov-Smirnoff, teste de Shapiro-Wilk, e a estatística Q de Royston e Wright (STASINOPOULOS *et al.*, 2008). A estatística Q é calculada para todos os parâmetros do modelo, fornecendo quatro valores: Q_1 , Q_2 , Q_3 e Q_4 , respectivamente para μ , σ , ν , τ . A distribuição dos z-escores ajustados é testada para a normalidade em cada faixa da intervalo da variável explicativa (idade) usando o teste de Shapiro-Wilk e os testes de D'Agostino modificados para assimetria e curtose. Isto fornece indicação de quais momentos da distribuição (média, assimetria ou curtose) estão sendo inadequadamente modelados em cada faixa da variável independente. Valores estatisticamente significantes de Q_1 , Q_2 , Q_3 e Q_4 indicam possível inadequação do respectivo parâmetro do modelo naquela faixa específica da variável independente. No GAMLSS, valores de $|Q| > 2$ devem ser considerados indicativo de inadequação do modelo (STASINOPOULOS *et al.*, 2008).

O *qq-plot* é aplicado quando se pressupõe que os dados seguem uma distribuição conhecida. O gráfico mostra os quantis da distribuição teórica no eixo horizontal e os quantis empíricos no eixo vertical. O *qq-plot* para dados com distribuição normal, denominado de gráfico de probabilidade normal, é o mais conhecido, porém pode ser adaptado para outras distribuições. A inspeção visual dos pontos distribuídos em torno da diagonal permite verificar desvios entre os quantis empíricos e os quantis teóricos. A adequação do modelo se verifica quando os pontos no gráfico se dispõem em linha reta coincidente com a diagonal ($y = x$) do gráfico, enquanto o afastamento da distribuição esperada é verificado quando este gráfico assume formas diferentes da linha reta. Por exemplo, se os pontos no gráfico assumem forma de “S”, indica distribuições com caudas muito curtas, isto é, distribuições cujos valores estão muito próximos da média (leptocurtose); em forma de “S” invertido: indicam distribuições com caudas muito longas e, portanto, presença de muitos valores extremos (platicurtose); forma de “J” e “J” invertido indica distribuições assimétricas, positivas e negativas, respectivamente. Na Figura 2 está representado um esquema para interpretação visual de um *qq-plot*.

experience of the researcher. However, one basic rule in the residue chart interpretation is that the appearance of a cloud of points with a clear shape instead of a random standard is a warning that the residues are not normally distributed. The next step is to look for help in the normality statistics.

Residue normality can also be verified by means of specific statistical tests, as for example, the Kolmogorov-Smirnoff test, Shapiro-Wilk test, and Royston and Wright's Q statistics (STASINOPOULOS *et al.*, 2008). Q -stats are calculated for all the model parameters, providing four values: Q_1 , Q_2 , Q_3 and Q_4 , respectively for μ , σ , ν , τ . The distribution of adjusted z-scores is tested for normality in each interval range of the explanatory variable (age) using the Shapiro-Wilk test and D'Agostino TST modified for asymmetry and kurtosis. This provides one indication of which distribution moments (mean, asymmetry or kurtosis) are being inadequately modeled in each range of the independent variable. Q_1 , Q_2 , Q_3 and Q_4 statistically significant values show possible inadequacy of the respective parameter of the model in that specific range of the independent variable. On GAMLSS, $|Q| > 2$ values must be considered as to be showing model inadequacy (STASINOPOULOS *et al.*, 2008).

QQ-plot is applied when it is assumed that the data follow one unknown distribution. The chart shows the theoretical distribution quantiles in the horizontal axis and the empirical quantiles in the vertical axis. *Qq-plot* for data with normal distribution, called normal probability chart, is the most well-known, however, it can be adapted for other distributions. The visual inspection of the distributed points around the diagonal allows seeing deviations between the empirical quantiles and the theoretical quantiles. The adequacy of the model can be verified when the points in the chart are displayed in a straight line which coincides with the chart diagonal ($y = x$), whereas the distancing of the expected distribution is verified when this chart takes up shapes other than the straight line. For example, if the points in the chart are in an “S” shape, this means distributions with too short tails, that is, distributions whose values are too close to the mean (leptokurtosis). A mirrored “S” shows distributions with too long tails and, consequently, presence of many extreme values (platykurtosis). “J” and mirrored “J” mean asymmetrical distributions, positive and negative, respectively. Figure 2 shows one framework for the visual interpretation of a *qq-plot*.

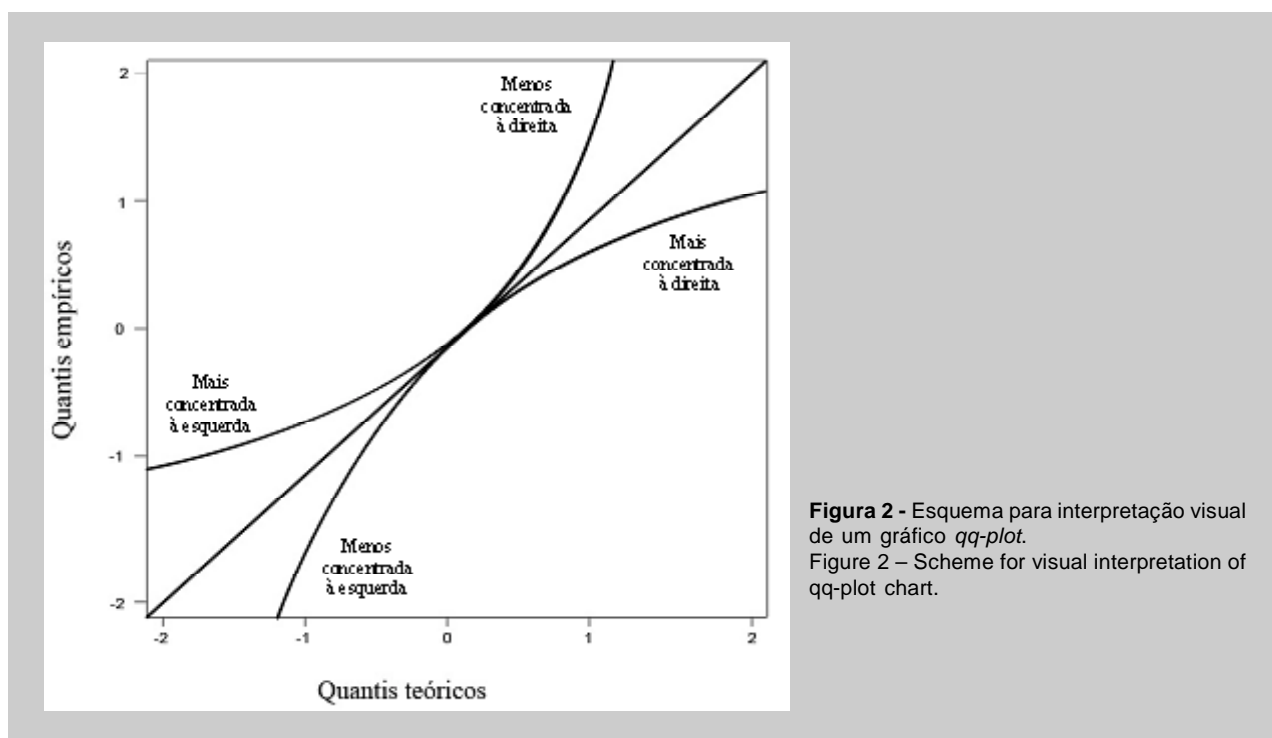


Figura 2 - Esquema para interpretação visual de um gráfico *qq-plot*.
Figure 2 – Scheme for visual interpretation of *qq-plot* chart.

Já na versão *detrended qq-plot* o eixo vertical do gráfico representa as diferenças entre os z-escores dos valores empíricos padronizados e os z-escores dos valores esperados sob a distribuição teórica, enquanto no eixo horizontal se encontram os z-escores dos valores teóricos. O *detrended qq-plot* enfatiza os desvios da linha reta teórica (Figura 3). O gráfico denominado *Worm Plot* (WP) de VAN BUUREN e FREDRIKS (2001) consiste de uma coleção de gráficos *detrended qq-plot*, cada um deles aplicado a um dos sucessivos intervalos obtidos pela partição do intervalo de variação da variável explicativa (Figura 4). O eixo vertical do WP retrata, para cada observação, a diferença entre a sua posição na distribuição teórica e na distribuição empírica. Os pontos no gráfico formam uma Figura com aspecto vermiforme, justificando a denominação inglesa de *worm plot*. Estão representados também os limites superior e inferior do intervalo de confiança de 95% que representa a variação amostral e delinea a região onde o “*verme*” deve se localizar (Figura 4). Se o modelo se ajusta bem aos dados em um determinado intervalo, os pontos devem cair dentro do intervalo de confiança e a Figura resultante apresenta uma forma plana. Uma Figura em forma de verme indica como os dados diferem dos valores preditos pelo modelo sob a distribuição teórica, sugerindo modificações no modelo. A principal vantagem do

In the *detrended qq-plot* version, the vertical axis of the chart represents the differences between the z-scores of the standardized empirical values and the expected value z-scores under the theoretical distribution, whereas the horizontal axis presents the theoretical value z-scores. The *detrended qq-plot* emphasizes the deviations of the theoretical straight line (Figure 3). The chart named *Worm Plot* (WP) by VAN BUUREN and FREDRIKS (2001) consists of a collection of *detrended qq-plot* charts, each one of them applied to one of the successive intervals obtained by the partition of the variation interval of the explanatory variable (Figure 4). The WP vertical axis depicts, for each observation, the difference between its position in the theoretical distribution and in the empirical distribution. The points in the chart form one wormlike figure, which justifies its name (*worm plot*). The upper and lower limits of the 95% confidence interval are also shown, which represents the sample variation and outlines the region where the “*worm*” must locate (Figure 4). If the model fits well to the data in one specific interval, the points must fall within the confidence interval and the resulting figure presents one flat form. One wormlike figure shows how the data differ from the values predicted by the model under the theoretical

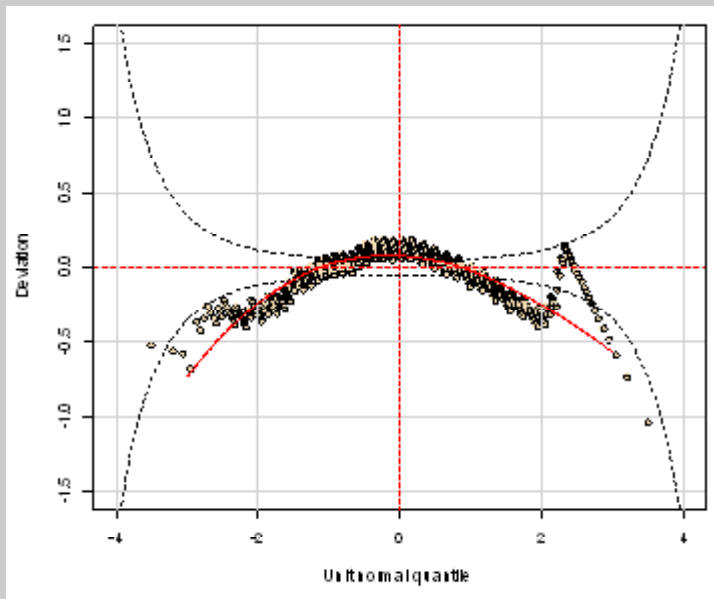


Figura 3 - Detrended QQ-plot do modelo ajustado como distribuição NORMAL. O gráfico mostra grande quantidade de pontos fora da região delimitada pelas bandas de confiança de 95%, indicando nítidas inadequações do modelo NORMAL.

Figure 3 – Detrended QQ-plot of the adjusted model with NORMAL distribution. The chart shows great quantity of points out of the region outlined by the 95% confidence strips, which shows clear inadequacies of the NORMAL model.

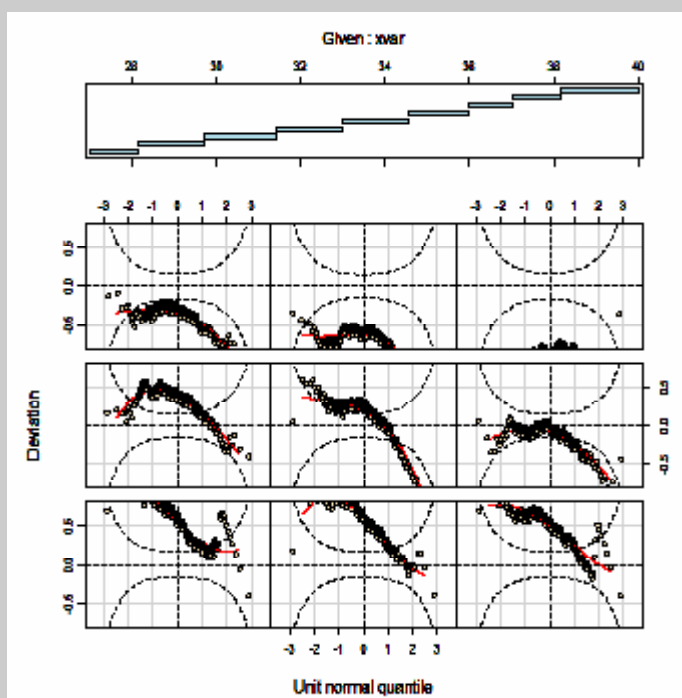


Figura 4 - Worm-plot do modelo ajustado com a distribuição NORMAL em nove faixas do intervalo da variável explicativa idade gestacional (IG). A Figura é composta por nove *detrended qq-plot* dispostos em sequência da esquerda para a direita e de cima para baixo, e que correspondem às nove faixas de idade gestacional (mostradas como degraus acima do *worm-plot* representando o intervalo de 27 a 40 semanas). Notar que quase todos os pontos dos gráficos estão plotados fora da região delimitada pelas bandas de confiança de 95% indicando falta de ajustamento do modelo normal.

Figure 4 – Worm-plot of the adjusted model with NORMAL distribution in nine areas of the Gestational Age (GA) explanatory variable interval. The Figure is comprised of nine detrended qq-plots organized in sequence from left to right and from top to bottom, which correspond to nine gestational age ranges (shown as steps above worm-plot representing the 27-40 week interval). Note that almost all points of the charts are plotted out of the region outlined by the 95% confidence strips, which indicates normal model's lack of fit.

WP é que pode indicar em que faixa do intervalo da variável explicativa o modelo não se ajusta adequadamente.

Como vários modelos podem ser simultaneamente bem ajustados aos dados e os pacotes estatís-

distribution, suggesting modifications in the model. WP's main advantage is that it can show in which zone of the explanatory variable interval the model will not fit adequately.

Since several models can simultaneously fit well

ticos permitem a construção de muitos modelos, é preciso então selecionar dentre os modelos que se mostrarem adequados o de melhor ajuste. A análise visual dos gráficos produzidos depende da subjetividade do pesquisador e nem sempre será suficiente. Por isso critérios objetivos devem ser utilizados na seleção de modelos. Um critério objetivo, porquanto é aritmético, é o desvio global (DG) ou deviance, que é definida como menos duas vezes o logaritmo da função de verossimilhança (na hipótese de erros com distribuição normal) calculado para os valores estimados dos coeficientes do modelo. No GAMLSS, este valor é fornecido pela função `deviance()`. Quanto menor a deviance, melhor ajuste fornece o modelo.

Uma abordagem para contornar o problema de modelos sobreajustados é utilizar os critérios de informação que levam em consideração o número de parâmetros e de graus de liberdade utilizados no modelo para penalizar os modelos mais complexos e evitar sobreajustamento aos dados em amostras de grandes tamanhos. Um critério geral para a seleção de modelos é o Critério de informação de Akaike generalizado (GAIC), que é uma estatística obtida pela adição ao deviance global (DG) de uma penalidade fixada p para cada grau de liberdade (gl) usado no modelo. Assim, $GAIC(p) = DG + p \cdot gl$. O critério de informação de Akaike (AIC) e o critério bayesiano de Schwarz (SBC) são casos especiais de $GAIC(p)$, em que a penalidade adotada é $p = 2$ e $p = \log(n)$, respectivamente (RIGBY e STASINOPOULOS, 2004). Esses critérios numéricos são bastante usados para discriminar modelos porque ponderam parcimônia (menos parâmetros) com melhor adequação (menores desvios). A regra é selecionar entre os modelos candidatos aquele que produzir o menor valor do AIC ou SBC. Menores valores para AIC e SBC indicam melhor ajuste, de modo que ambos os critérios permitem comparação entre modelos não aninhados e penalizam aqueles com maior número de parâmetros.

Aplicação do método GAMLSS

A aplicação do método GAMLSS é demonstrada neste artigo através da construção de curvas de referência para o índice de resistência da artéria cerebral média de fetos saudáveis entre 27 e 39 semanas de gestação, utilizando os dados de um estudo de PAIVA *et al.* (2008) que forneceu um banco de dados com duas variáveis contínuas: a variável explicativa idade gestacional (IG), manipulada continuamente e a variável dependente índice de resistência na artéria cerebral média (IR). A amostra era constituída de 2.236 medidas.

to the data, and the statistical packages allow the construction of many models, it is necessary to choose the model that fits best from those that prove to be adequate. The visual analysis of the charts produced depends on the researcher's subjectivity and will not always be enough. For this, objective criteria must be used in the selection of the models. One objective criterion, while arithmetic, is the Global Deviance (GD), which is defined as minus two times the logarithm of the function of verisimilitude (in case of errors with normal distribution) calculated for the estimated values of the model's coefficients. In GAMLSS, this value is provided by the `deviance()` function. The littler the deviance, the better fit the model will supply.

One approach to overcome the problem of overfitting models is to use the criteria of information that take into consideration the number of parameters and degrees of freedom used in the model to penalize the more complex models and avoid overfitting to the data in big-size samples. One general criterion for the selection of models is the Generalized Akaike Information Criterion (GAIC), which is one statistics obtained by the addition to the Global Deviance (GD) of one fixed penalty p for each degree of freedom (df) used in the model. Then, $GAIC(p) = GD + p \cdot df$. Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) are special cases of $GAIC(p)$, in that the adopted penalty is $p = 2$ and $p = \log(n)$, respectively (RIGBY and STASINOPOULOS, 2004). These numeric criteria are very much used to differentiate models because they ponder parsimony (fewer parameters) with better adequacy (littler deviations). The rule is to select from the candidate models the one which produced the lowest value of AIC or SBC. Lower AIC and SBC values show better fit, in a way that both criteria allow comparison between non-aligned models and penalize those with more parameters.

Application of the GAMLSS method

The application of the GAMLSS method is demonstrated in this article through the construction of reference curves for the resistance index of the middle brain artery of healthy fetuses between 27 and 39 weeks of gestation, using the data from the study by PAIVA *et al.* (2008) which supplied a database with two continuous variables: The gestational age explanatory variable (GA), manipulated continuously, and the resistance index dependent variable of the middle brain artery (RI). The sample was comprised of 2,236 measures.

Para elaborar o modelo estatístico e construir as curvas de referência, é necessária uma análise do gráfico de dispersão da variável IR em função da IG (Figura 5) para responder às seguintes questões preliminares: a) como é que a média do IR depende da IG? b) como é que a variância do IR depende de sua média? c) qual é a distribuição do IR em dada IG? d) os parâmetros de escala e forma da distribuição do IR dependem da IG?

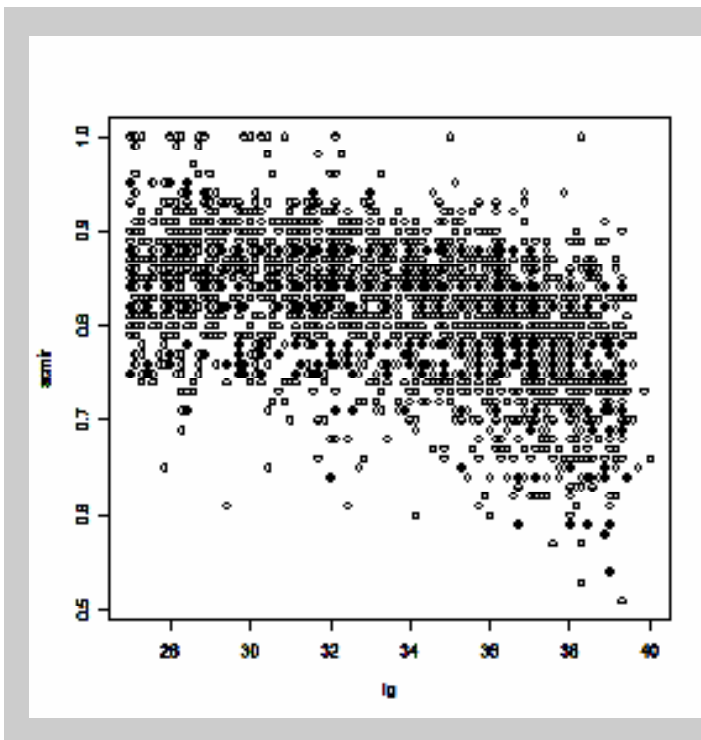


Figura 5 - Gráfico de dispersão de 2.236 valores da variável índice de resistência da artéria cerebral média (IR) em função da idade gestacional (IG) de 27 a 39 semanas (dados de Paiva *et al.*, 2008).

Figure 5 – Chart of the dispersion of 2,236 values of the resistance index variable of the middle brain artery (RI) in function of Gestational age (GA) from 27 to 39 weeks (data by Paiva *et al.*, 2008)

A nuvem de pontos revela uma possível correlação negativa entre as variáveis IG e IR e que a variabilidade do IR em um dado valor de IG não parece ser constante ao longo de todo o intervalo da variável IG. Além disso, o histograma da distribuição univariada de IR (Figura 6a) mostra que a distribuição amostral é assimétrica e levemente leptocúrtica. Para a terceira questão, utiliza-se um subconjunto dos dados para verificar como a variável IR se distribui numa faixa específica de IG. Neste exemplo, escolheu-se arbitrariamente a faixa entre 29 e 30 semanas. O histograma desse subconjunto (Figura 6b) revela uma assimetria à esquerda (coeficiente de assimetria = 0,392) e uma curtose maior que três (coeficiente de curtose = 4,055). Percebe-se então que a distribuição da variável IR nesse subconjunto é moderadamente assimétrica e leptocúrtica, pois apresenta coeficiente de assimetria diferente de zero e curtose maior do que três. Esta característica também é

The cloud of points reveals one possible negative correlation between the GA and RI variables and that the RI variability at a given GA value does not seem to be constant throughout the GA variable interval. Besides, the RI univariated distribution histogram (Figure 6a) shows that the sample distribution is asymmetric and slightly leptokurtic. For the third question, a subgroup of data is used to verify how the RI variable distributes in one GA specific zone. In this example, the range between 29 and 30 weeks was randomly chosen. This subset histogram (Figure 6b) reveals one asymmetry to the left (asymmetry coefficient = 0,392) and a kurtosis over 3 (kurtosis coefficient = 4,055).

It can be seen, then, that the RI variable distribution in this subgroup is moderately asymmetric and leptokurtic, for it shows an asymmetry coefficient different from zero and kurtosis over three. This characteristic is also presented by the RI univariated

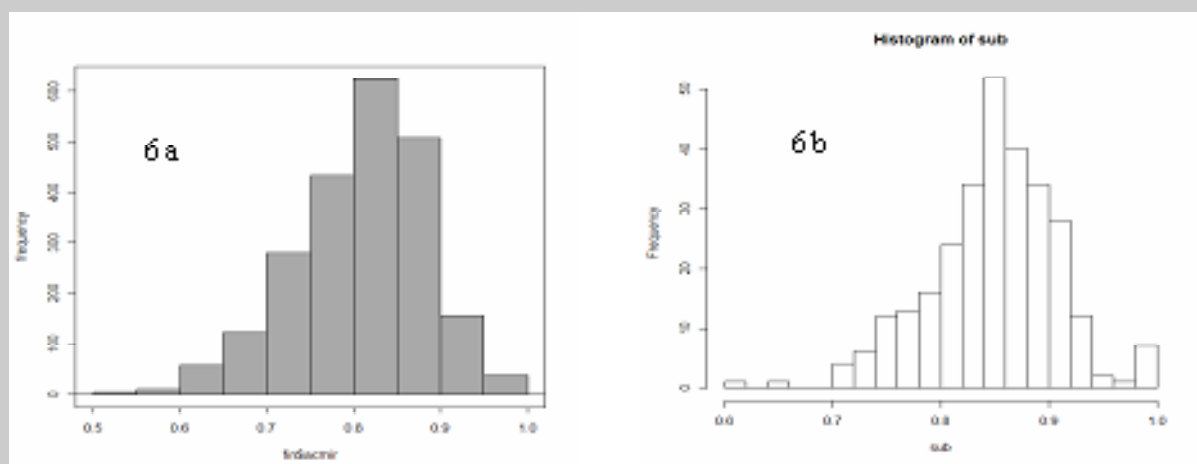


Figura 6 - (a) Histograma da distribuição univariada de 2.236 valores da variável índice de resistência da artéria cerebral média (IR) de 27 a 39 semanas.

(b) Histograma da distribuição univariada de um subconjunto de valores da variável IR na faixa de 29 a 30 semanas (dados de Paiva *et al.*, 2008).

Figure 6 – (a) Histogram of the univariad distribution of 2,236 values of resistance index variable of middle brain artery (RI) from 27 to 39 weeks. (b) Histogram of the univariad distribution of a subgroup of RI variable values in the range of 29 to 30 weeks (data by Paiva *et al.*, 2008)

apresentada pela distribuição univariada de IR em todo o intervalo da variável explicativa IG, conforme se observa pelo histograma mostrado na Figura 6a.

Inicialmente, tenta-se ajustar o modelo admitindo-se a distribuição normal e utilizam-se as ferramentas gráficas de diagnóstico para avaliar a qualidade do ajuste do modelo. A função `plot()` do GAMLSS gera quatro gráficos que são apresentados na Figura 1. Os dois gráficos superiores são os resíduos normalizados plotados contra os valores ajustados (Figura 1a) e contra um índice de ordenação dos dados (Figura 1b). O aspecto da distribuição dos resíduos parece randômico, mas notam-se muitos pontos fora do intervalo $[-2, +2]$. Se não houver um significativo afastamento dos pressupostos de normalidade dos resíduos, estes gráficos colaboram pouco com o diagnóstico do modelo. No entanto, o gráfico qq-plot (Figura 1d) é bastante informativo. Nota-se claramente o afastamento dos pontos nas extremidades do gráfico, confirmando que a distribuição de IR é menos concentrada à esquerda (Figura 2).

A versão *detrended qq-plot* ou *worm plot* (WP)

distribution throughout the GA explanatory variable, as can be seen by the histogram shown in Figure 6a.

In the beginning, adjusting the model is attempted assuming the normal distribution, and the diagnostic graphic tools are used to evaluate the quality of the model fit. The GAMLSS plot function() generates four charts that are presented in Figure 1. The two upper charts are the normalized residue plotted against the fit values (Figure 1a) and against one data ordination index (Figure 1b). The distribution aspect of the residues seems random, but one can see many points out of the interval $[-2, +2]$. If there is not one significant distancing of the residue normality presuppositions, these charts contribute little to the model diagnosis. However, the qq-plot chart (Figure 1d) is very informative. The distancing of the points at the edge of the chart can be clearly seen, confirming that the RI distribution is less concentrated to the left (Figure 2).

The *detrended qq-plot* version or *worm plot* (WP) by VAN BUUREN and FREDRIKS (2001) is much

de VAN BUUREN e FREDRIKS (2001) é muito mais fácil de utilizar, pois é mais intuitivo e vem com os limites do intervalo de confiança de 95% que delimita a região de adequação do modelo. Esta ferramenta pode ser utilizada como um único gráfico expressando todo o intervalo da variável explicativa (Figura 3) ou como uma coleção de gráficos *detrended qq-plot* em várias faixas do intervalo (Figura 4). Através da análise visual das Figuras 3 e 4, percebe-se a inadequação do modelo NORMAL em todas as faixas de IG, pois quase todos os pontos estão plotados fora da região delimitada pelo intervalo de confiança de 95%.

A melhor ferramenta para testar se os valores dos resíduos seguem distribuição normal ao longo de todo o intervalo da IG é a estatística Q de ROYSTON e WRIGHT (1997), pois é um critério objetivo. No GAMLSS é calculada pela função Q.stats. O número de intervalos é definido pelo pesquisador. Nesta demonstração optou-se por particionar a variável IG em 11 intervalos. A estatística K2 do teste de D'Agostino-Pearson para normalidade dos resíduos segue uma distribuição qui-quadrado com 2 graus de liberdade. A hipótese alternativa do teste é "a distribuição dos resíduos é não normal". Valores significantes de Q_1 , Q_2 , Q_3 e Q_4 indicam inadequação do modelo respectivamente para os parâmetros μ , σ , ν e τ . Como uma regra prática para interpretação desse teste, RIGBY e STASINOPOULOS (2008) sugerem que valores de $|Z|$ superiores a 2 devem ser considerados como indicativos de significativas inadequações no modelo. Nesta aplicação, verifica-se que o módulo de Z é maior do que 2 em várias faixas de idade, indicando que o modelo com distribuição normal é inadequado. Isto é confirmado pelo p-valor altamente significativo da estatística Q (Tabela 1).

Após o diagnóstico de que o modelo NORMAL é inadequado, e que a distribuição dos dados é assimétrica e leptocúrtica, o passo agora é construir modelos com as várias famílias de distribuições disponíveis contínuas no GAMLSS, com um, dois, três ou quatro parâmetros. As distribuições Box-Cox t (BCT), Johnson's SU (JSU), skew exponencial Power (SEP), Box-Cox power exponencial (BCPE) são capazes de modelar tanto assimetria quanto curtose. A distribuição t family (TF) só é capaz de modelar curtose, enquanto a Box-Cox Cole Green (BCCG) apenas modela assimetria

easier to use, for it is more intuitive and it comes with the 95% confidence interval limits, which outlines the model's region of fit. This tool can be used as one single chart expressing all the explanatory variable interval (Figure 3) or as a collection of *detrended qq-plot* charts in many areas of the interval (Figure 4). Through the visual analysis of Figures 3 and 4, the inadequacy of the NORMAL model in all the GA areas can be noticed, for almost all the points are plotted out of the outlined region by the 95% confidence interval.

The best tool to test if the residue values follow the normal distribution throughout the GA interval is the Q-stats by ROYSTON and WRIGHT (1997), for it is one objective criterion. In GAMLSS, it is calculated by the Q-stats function. The number of intervals is defined by the researcher. In this demonstration, we chose to break the GA variable into 11 intervals. The K2 statistics of the D'Agostino-Pearson test for the residue normality follows one chi-squared distribution with 2 degrees of freedom. The alternative hypothesis of the test is "the residue distribution is non normal". Significant Q_1 , Q_2 , Q_3 and Q_4 values show inadequacy of the model for the μ , σ , ν and τ parameters respectively. As a practical rule for the interpretation of this test, RIGBY and STASINOPOULOS (2008) suggest that $|Z|$ values over 2 must be considered as to indicate significant inadequacies in the model. In this application, it can be seen that the Z module is over 2 in many age ranges, showing that the model with normal distribution is inadequate. This is confirmed by Q-stats highly significant p-value (Table 1).

After the diagnosis that the NORMAL model is inadequate, and that the data distribution is asymmetric and leptokurtic, the step now is to build models with the several continuous distribution families available at GAMLSS, with one, two, three or four parameters. The Box-Cox t distributions (BCT), Johnson's SU (JSU), skew exponential Power (SEP), Box-Cox power exponential (BCPE) are capable of modeling both asymmetry and kurtosis. The t family distribution (TF) is only capable of modeling kurtosis, whereas Box-Cox

Tabela 1 - Valores da estatística Q de Royston e Wright (2000) para o modelo ajustado com distribuição normal da variável índice de resistência da artéria cerebral média de fetos normais em função da idade gestacional (IG) entre 27 e 39 semanas.

Table 1 – Q-stats values by Royston and Wright (2000) for the adjusted model with normal distribution of the resistance index variable of the middle brain artery of normal fetuses in function of Gestational age (GA) between 27 and 39 weeks.

IG(semanas)	Z1	Z2	Z3	Z4	Agostino K2	N
27 a 28	8,02	-5,30	1,48	2,18	6,96	223
28 a 29,1	8,78	-5,05	-1,17	1,11	2,59	210
29,1 a 30,4	7,37	-3,36	-2,57	2,57	13,20	181
30,4 a 31,8	4,92	-4,95	-1,73	0,49	3,24	203
31,8 a 33,1	4,57	-2,96	-3,86	1,84	18,27	217
33,1 a 34,2	1,73	-3,99	-3,11	0,30	9,76	188
34,2 a 35,5	-0,40	-3,26	-1,39	-0,55	2,25	217
35,5 a 36,4	-4,73	-1,92	-2,87	-0,36	8,35	202
36,4 a 37,4	-7,83	-1,43	-1,76	-0,95	4,01	206
37,4 a 38,2	-10,45	-1,08	-2,06	1,64	6,94	207
38,2 a 40	-12,71	0,04	-1,51	-0,08	2,29	182
TOTAL Q stats	598,58	131,58	57,50	20,35	77,85	2236
df for Q stats	10	10	11	11	22	
p-val for Q stats	3,54e-122	2,22e-23	2,69e-08	0,04	3,64e-08	

(STASINOPOULOS E RIGBY, 2008). Foram construídos modelos para várias distribuições, porém o alisador não paramétrico *splines* cúbicas foi utilizado em todos os modelos testados. Entre os vários termos aditivos disponíveis no GAMLSS (cubic *splines*, varying coefficient, penalized *splines*, loess, fractional polynomials), optou-se por utilizar *splines* cúbicas porque este alisador mostrou-se mais adequado no estudo da OMS (WHO, 2006).

Para selecionar o modelo de melhor ajuste, foi utilizado o critério de informação de Akaike. A regra é selecionar entre os modelos candidatos aquele que produzir o menor valor do AIC ou SBC. A Tabela 2 mostra os valores do AIC e do SBC para os vários modelos testados. Pelo AIC o modelo selecionado foi aquele com a distribuição Johnson's SU (JSU). Se o critério adotado fosse o SBC, o melhor modelo seria o Box-Cox Cole e Green (também conhecido como método LMS). Após ter selecionado o modelo, faz-se um diagnóstico da qualidade do ajuste com as ferramentas acima citadas.

Os gráficos dos resíduos normalizados bem como o qq-plot (Figura 7) sinalizam para melhor qualidade do ajustamento do modelo selecionado em comparação com o modelo normal (Figura 1). Entretanto, a ferramenta *worm plot* é muito mais elucidativa e mais fácil de interpretar. Os gráficos das nove faixas de IG mostram o

Cole Green (BCCG) only models asymmetry (STASINOPOULOS and RIGBY, 2008). Models were built for several distributions; however, the cubic splines non-parametric smoother was used in all the tested models. Among the many available additive terms at GAMLSS (cubic *splines*, varying coefficient, penalized *splines*, loess, fractional polynomials), cubic splines were chosen to be used because this smoother has shown to be more adequate in the WHO study (WHO, 2006).

To choose the best-fitting model, the Akaike Information Criterion was used. The rule is to choose from the candidate models the one which produces the lowest AIC or SBC value. Table 2 shows AIC and SBC values for the many tested models. By the AIC, the chosen model was the one with the Johnson's SU distribution (JSU). If the adopted criterion were the SBC, the best model would be Box-Cox Cole and Green (also known as LMS method). After the choice of the model, a fitting quality diagnosis is carried out with the above mentioned tools.

The normalized residue charts as well as the qq-plot (Figure 7) show a better fit quality of the chosen model in comparison with the normal (Figure 1). However, the *worm plot* tool is much more clarifying and easier to interpret. The charts of the nine GA areas

Tabela 2 - Valores do critério de informação de Akaike (AIC) e do critério bayseano de Schwarz (SBC) para modelos ajustados com várias distribuições contínuas e alisador não-paramétrico do tipo *splines* cúbicas.

Table 2 – Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) values for adjusted models with several continuous distributions and cubic splines-like non parametric smoother.

Distribuição	Graus de liberdade	AIC	SBC
Johnson's SU	12	-5921,7*	-5853,2
Box-Cox Cole e Green	11	-5918,4	-5855,6**
Box-Cox t	12	-5918,9	-5850,4
Box-Cox power exponencial	12	-5918,5	-5850,0
Skew t type 1	12	-5907,1	-5838,5
Skew exponencial power	12	-5892,0	-5823,5
T family	11	-5886,3	-5823,5
Power exponencial	11	-5885,3	-5822,5
Normal	10	-5881,0	-5823,9
Weibull	10	-5836,8	-5779,7
Gamma	10	-5830,0	-5772,9
Log normal	10	-5796,7	-5739,6
Inverse Gaussian	10	-5795,8	-5738,7
Reverse Gumbell	10	-5288,1	-5231,0

* menor valor do AIC seleciona o modelo Johnson's SU

*lower AIC value selects Johnson's SU model

** menor valor do SBC seleciona o modelo Box-Cox Cole e Green

** lower SBC value selects Box-Cox Cole and Green model

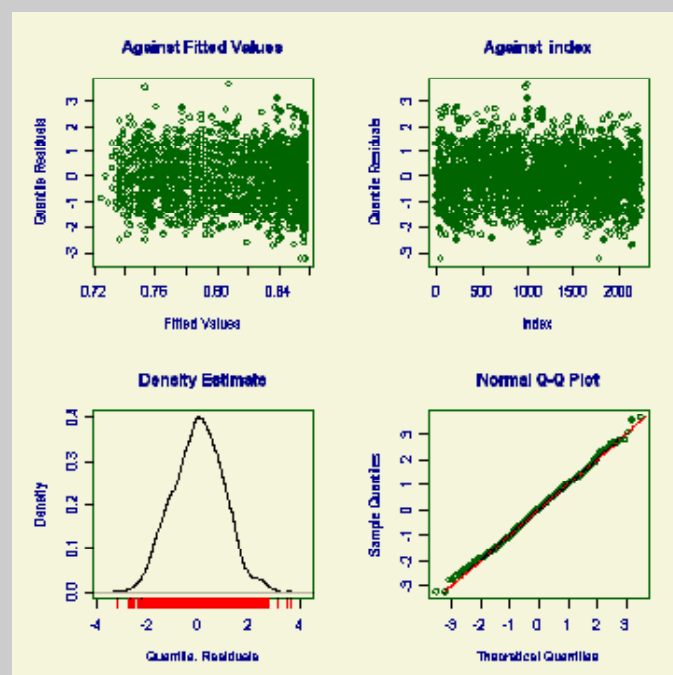
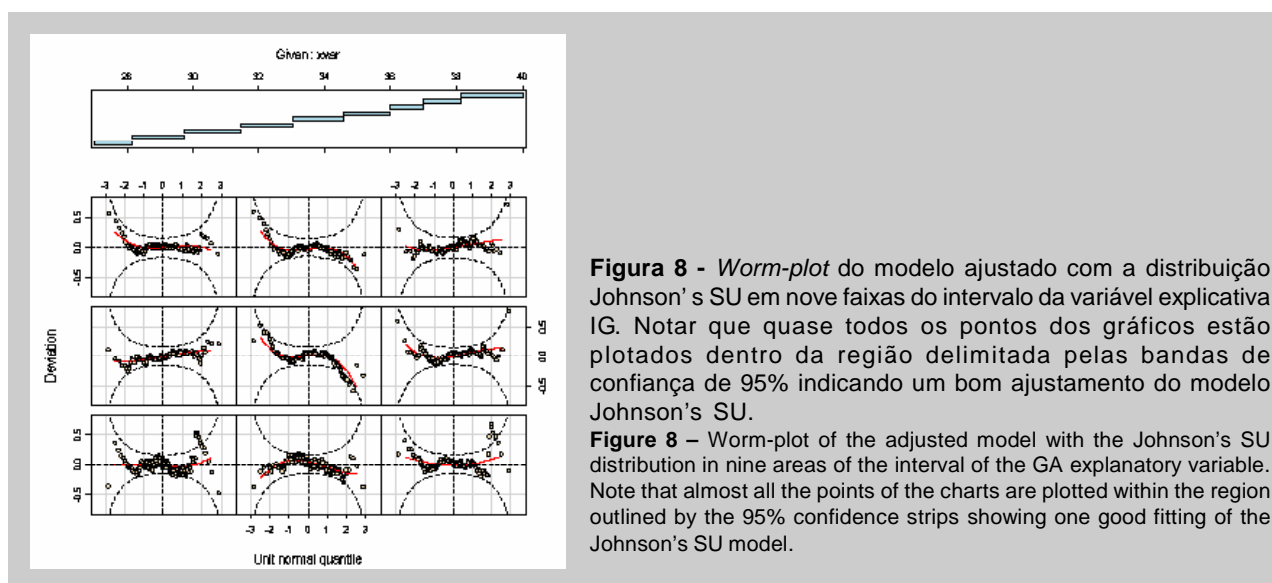


Figura 7 - Diagnóstico do modelo ajustado com a distribuição Johnson's SU. O gráfico dos resíduos normalizados e o qq-plot mostram a melhor ajustamento do modelo JSU em comparação com o modelo ajustado pela distribuição Normal (ver Figura 1).

Figure 7 – Diagnosis of the adjusted model with Johnson's SU distribution. The normalized residue and qq-plot chart show better fit of the JSU model in comparison with the model adjusted by normal distribution (see Figure 1).

excelente ajuste do modelo JSU em todas as faixas de IG, pois quase todos os pontos estão plotados dentro da região delimitada pelo intervalo de confiança de 95% (Figura 8). A comparação com a Figura 4 permite verificar a diferença significativa na qualidade do ajustamento

show the excellent fit of the JSU model in all GA ranges, for almost all the points are plotted within the outlined region by the 95% confidence interval (Figure 8). The comparison with Figure 4 allows the verification of the significant difference in the quality of the adjustment



produzido pelo modelo JSU em contraste com o modelo NORMAL. É interessante também demonstrar que os valores dos resíduos seguem distribuição normal ao longo de todo o intervalo da IG através da estatística Q. Na Tabela 3 verifica-se que o módulo de Z é menor do que 2 em todas as faixas de IG (exceto o valor para o parâmetro tau de uma única faixa), indicando que o modelo JSU é adequado. Isto é confirmado pelo p-valor não significativo da estatística Q.

produced by the JSU model in contrast with the NORMAL model. It is also interesting to demonstrate that the residue values follow a normal distribution throughout the GA interval by the Q-stats. Table 3 shows that the Z model is below 2 in all GA ranges (except the value for the tau parameter in one single range), which shows that the JSU model is adequate. This is confirmed by the non significant p-value in the Q-stats.

Tabela 3 - Valores da estatística Q de Royston e Wright (2000) para o modelo ajustado com distribuição Johnson's SU da variável índice de resistência da artéria cerebral média de fetos normais em função da idade gestacional (IG) entre 27 e 39 semanas.

IG(semanas)	Z1	Z2	Z3	Z4	AgostinoK2	N
27,0 a 28,1	-0,03	0,21	0,70	0,88	1,27	251
28,1 a 29,2	0,18	-0,71	-0,69	0,60	0,83	204
29,2 a 31	0,10	1,17	-0,12	0,94	0,89	228
31,0 a 32,2	0,07	-0,12	0,70	0,41	0,66	210
32,2 a 33,8	-0,43	-0,17	-0,24	-1,85	3,49	226
33,8 a 35,1	0,71	-0,19	0,85	0,20	0,77	235
35,1 a 36,1	-0,39	0,51	0,09	2,44	5,98	226
36,1 a 37,1	0,01	-0,76	0,34	0,13	0,13	211
37,1 a 38,1	-0,35	-0,69	-0,13	-1,77	3,17	232
38,1 a 40	0,13	0,78	0,45	0,01	0,20	213
TOTAL Q stats	1,04	3,94	2,60	14,79	17,39	2236
df for Q stats	2	7	5	9	14	-
p-val for Q stats	0,5952	0,7866	0,7611	0,09670	0,2360	-

O GAMLSS ainda possui outras ferramentas que permitem um maior refinamento dos parâmetros do modelo que não serão abordados neste trabalho. O passo seguinte é a geração do gráfico com as curvas centílicas e a confecção de uma tabela com os valores de predição dos centis mais utilizados nas aplicações

GAMLSS still possesses other tools that allow a better refinement of the model parameters which are not addressed in this paper. The next step is the generation of a chart with the centile curves and the confection of one table with the prediction values for the most used centiles in healthcare applications (2,5%, 5%, 10%, 50%,

na área de saúde (2,5%, 5%, 10%, 50%, 75%, 90%, 95%, 97,5%), etapa bastante facilitada no GAMLSS que dispõe de funções específicas para esta tarefa. Nessa aplicação exemplo foram geradas as curvas e a tabela com os centis de referência 2,5; 5; 50; 95 e 97,5 do IR da artéria cerebral média de fetos entre 27 e 39 semanas (Figura 9 e Tabela 4). O modelo obtido indica que o IR da artéria cerebral média de fetos sadios decresce com a idade gestacional entre 27 e 39 semanas. Uma abordagem para verificar a qualidade do ajustamento, após o cálculo dos centis de referência, é observar a percentagem dos valores que ficaram abaixo dos centis observados. No modelo JSU a frequência de casos abaixo do centil 2,5 é 2,0%; abaixo do centil 5 é 5,08%; abaixo do centil 50 é 48,82%; abaixo do centil 95 é 95,63% e abaixo do centil 97,5 é 97,41%.

75%, 90%, 95%, 97,5%), a step that is greatly eased by GAMLSS, which has specific functions for this task. In this example application, the curves and the table with the 2,5; 5; 50; 95 and 97,5 RI reference centiles of the middle brain artery of fetuses between 27 and 39 weeks were generated (Figure 9 and Table 4). The obtained model shows that the RI of the middle brain artery of healthy fetuses decreases with gestational age between 27 and 39 weeks. One approach to verify fit quality, after the calculation of reference centiles, is to observe the percentage of values that stayed below the observed centiles. In the JSU model, the frequency of cases below the 2,5 centile is 2,0%; below the 5 centile is 5,08%; below the 50 centile is 48,82%; below the 95 centile is 95,63% and below the 97,5 centile is 97,41%.

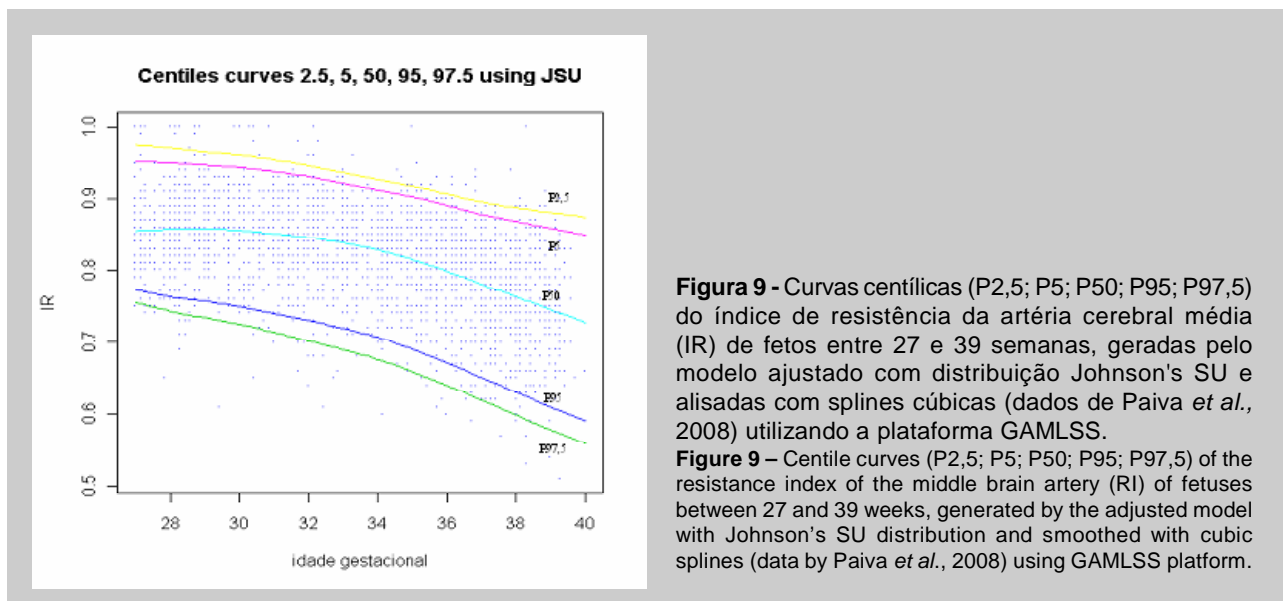


Figura 9 – Curvas centílicas (P2,5; P5; P50; P95; P97,5) do índice de resistência da artéria cerebral média (IR) de fetos entre 27 e 39 semanas, geradas pelo modelo ajustado com distribuição Johnson's SU e alisadas com splines cúbicas (dados de Paiva *et al.*, 2008) utilizando a plataforma GAMLSS.

Figure 9 – Centile curves (P2,5; P5; P50; P95; P97,5) of the resistance index of the middle brain artery (RI) of fetuses between 27 and 39 weeks, generated by the adjusted model with Johnson's SU distribution and smoothed with cubic splines (data by Paiva *et al.*, 2008) using GAMLSS platform.

Tabela 4 - Valores de predição dos percentis 2,5; 5; 50; 95 e 97,5 do índice de resistência da artéria cerebral média de fetos normais em função da idade gestacional (IG) de 27 a 39 semanas.

Table 4 – Resistance index 2,5; 5; 50; 95 and 97,5 percentiles prediction values of middle brain artery of normal fetuses in function of gestational age (GA) for 27 to 39 weeks.

IG	P2,5	P5	P50	P95	P97,5
27	0,75	0,77	0,86	0,95	0,97
28	0,74	0,76	0,86	0,95	0,97
29	0,74	0,76	0,86	0,95	0,97
30	0,73	0,75	0,86	0,95	0,96
31	0,72	0,74	0,85	0,94	0,96
32	0,70	0,73	0,84	0,93	0,95
33	0,69	0,72	0,84	0,92	0,94
34	0,68	0,71	0,83	0,92	0,93
35	0,67	0,70	0,81	0,91	0,92
36	0,65	0,68	0,80	0,89	0,91
37	0,63	0,65	0,78	0,88	0,90
38	0,60	0,63	0,76	0,87	0,89
39	0,58	0,61	0,74	0,86	0,88

Demonstrou-se com essa aplicação com dados empíricos que o método GAMLSS é bastante flexível e eficiente na modelagem estatística de curvas de referência. Nele estão disponíveis várias famílias de distribuição para elaboração do modelo matemático, além de ferramentas de diagnóstico e seleção do modelo com melhor ajuste, e também funções específicas para a construção das curvas de referência e cálculo facilitado dos percentis ou z-escores. Além disso, o método é acessível à comunidade científica e exige recursos computacionais mínimos, podendo ser amplamente utilizado em qualquer microcomputador com configuração básica. As evidências científicas reforçam que o método GAMLSS é uma importante ferramenta de modelagem estatística na área da saúde e que em futuras pesquisas provavelmente será o estado da arte na construção de novas curvas de referência ou na validação de curvas já construídas.

With this application with empirical data, it was shown that the GAMLSS method is fairly flexible and efficient in the statistical modeling of reference curves. In it, many distribution families for the design of a mathematical model are available, besides diagnostic tools and the selection of the best-fitting model, as well as specific functions for the construction of reference curves and eased calculations of the percentiles or z-scores. Moreover, the method is accessible to the scientific community and demands minimal computer resources, and it can be run on any microcomputer with basic configuration. The scientific evidence reinforces that the GAMLSS method is one important modeling statistical tool in healthcare and that in future researches it will be the state of the art in the construction of new reference curves or in the validation of curves already built.

REFERÊNCIAS

References

- AKANTZILIOTOU K, RIGBY RA, STASINOPOULOS DM. The R implementation of Generalized Additive Models for Location, Scale and Shape in Statistical modeling. in Anais do Congresso: Proceedings of the 17th International Workshop on statistical modelling, 2002. Chania, Greece, ed: Stasinopoulos M and Touloumi G, 2002.
- BRIASSOULIS H. Analysis of Land Use Change: Theoretical and Modeling Approaches. In: The Web Book of Regional Science. Disponível em: [http://www.rri.wvu.edu/WebBook/Briassoulis/Chapter4\(Models1\).htm](http://www.rri.wvu.edu/WebBook/Briassoulis/Chapter4(Models1).htm). Acesso em: 20/10/2009.
- BORGHI E, de ONIS M, GARZA C, VAN DEN BROECK J, FRONGILLO EA, GRUMMER-STRAWN L, et al. WHO Multicentre Growth Reference Study Group. Construction of the World Health Organization child growth standards: selection of methods for attained growth curves. *Stat Med* 25(s.n):247-265, 2006.
- COLE TJ, GREEN PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med* 11(s.n):1305-1319, 1992.
- DIBLEY MJ, GOLDSBY J, STAEHLING N, TROWBRIDGE FL. Development of normalized curves for the international growth reference: historical and technical considerations. *Am J Clin Nutr* 46(s.n):736-748, 1987.
- DUNN KP, SMYTH GK. Randomized quantile residuals. *J Comput Graph Stat* 5(1-10): 236-244, 1996.
- HASTIE TJ, TIBSHIRANI RJ. Generalized Additive Models: Some Applications. *J Am Stat Assoc* 82(398):371-386, 1987.
- MASSAD, E. Modelos Matemáticos em Biomedicina. In: Massad E, Menezes RX; Silveira PSP, Ortega NRS (Org.). Métodos Quantitativos em Medicina. Barueri: Manole, 2004.
- ONIS M, ONYANGO AW, BORGHI E, SIYAM A, NISHIDA C, SIEKMANN J et al. . Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 85(9):660-667, 2007.
- PAIVA CSM, LIMA MD, FREIRE DMC. Curvas percentilares da relação umbilico-cerebral em função da idade gestacional: uma aplicação do método LMS e alisamento com *splines* cúbicas. In: XXV CONGRESSO NORDESTINO DE GINECOLOGIA E OBSTETRÍCIA, 2008, João Pessoa: Idéia, 2008.
- RIGBY RA, STASINOPOULOS DM. The GAMLSS project: a flexible approach to statistical modelling, in: New Trends in Statistical Modelling: Proceedings of the 16th International Workshop on Statistical Modelling, Odense, 2001. Denmark, Klein B and Korsholm L, 2001.
- RIGBY RA, STASINOPOULOS DM. Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Stat Med* 23(s.n):3053-3076, 2004.
- RIGBY RA, STASINOPOULOS DM. Generalized Additive Models for Location, Scale and Shape, (with discussion). *Appl Stat* 54(s.n):507-554, 2005.
- RIGBY RA, STASINOPOULOS DM. Instructions on how to use the gamlss package in R. Second Edition. Jun 2008. Disponível em: <http://www.gamlss.com/>. Acesso em 03/11/2008.
- STASINOPOULOS DM, RIGBY RA. Generalized additive models for location scale and shape (GAMLSS) in R. *J Stat Soft* 23(7): 1-46, 2007.
- VAN BUUREN S, FREDRIKS M. Worm plot. A simple diagnostic device for modeling growth reference curves. *Stat Med* 20(s.n):1259-1277, 2001.
- WHO. Multicentre Growth Reference Study Group. 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, 2006.
- WRIGHT EM, ROYSTON PA. A comparison of statistical methods for age-related reference intervals. *J Royal Stat Soc* 160:47-69, 1997.

CORRESPONDÊNCIA

Correspondência

Cláudio Sérgio Medeiros Paiva
Rua Anselmo Gomes da Silva, 51. Bessa.
58.000-000 João Pessoa - Paraíba - Brasil

E-mail:

claudiosmp@terra.com.br
rebrasa@ccs.ufpb.br