Universidad Santo Tomas
Primer Claustro Universitario de Colombia

Comunicaciones en Estadística
Junio 2017, Vol. 10, No. 1, pp. 113–128

A Calibration Function Built From Change Points: a Review

Una función de calibración construida a partir de puntos de cambio: revisión

Ehidy Karime Garcia Cruz[a]
ekgarciac@unal.edu.co

Juan Carlos Correa M.[b]
jccorrea@unal.edu.co

Juan Carlos Salazar U.[c]
jcsalaza@unal.edu.co

Resumen

El problema de calibración no es reciente. Los trabajos en este tema fueron presentados inicialmente por Krutchkoff en la época de los 60, bajo un enfoque paramétrico y han sido ampliamente estudiados por otros autores desde diferentes perspectivas. Las investigaciones recientes respecto al punto de cambio han considerado supuestos adicionales y estimación usando modelos lineales mixtos. Se presenta una revisión exhaustiva de los problemas de calibración y punto de cambio. Adicionalmente, se puede observar que la vinculación de estos bajo el enfoque de modelos para datos longitudnales no ha sido trabajado.

***Palabras clave*:** Calibración, modelos mixtos, punto de cambio.

Abstract

Calibration is not a new problem, early in the 1960, it was worked by Krutchkoff under a parametric approach. Then, this idea has been widely studied and several approaches have been presented by different authors. On the other hand, latest proposals on change points consider some additional assumptions and worked based on a linear mixed models approach. We present an extensive review about this topics to show that both problems: calibration and change point has not been worked jointly yet under a longitudinal setting.

***Keywords*:** Calibration, linear mixed models, change point.

[a]Ph.D.(C) in Statistics, Universidad Nacional de Colombia, sede Medellín, Colombia
[b]Associated professor, Universidad Nacional de Colombia, sede Medellín, Colombia
[c]Associated professor, Universidad Nacional de Colombia, sede Medellín, Colombia

1. Introduction

In general, if we have two variables and we want to explore the nature of the relationship between them, we can consider a model given by

$$\mathbf{Y} = f(\mathbf{X};\boldsymbol{\beta}) + \boldsymbol{e},$$

where $\boldsymbol{e}_i \sim \mathbf{N}(\mathbf{0}, \sigma^2)$ and $f(\mathbf{X};\boldsymbol{\beta})$ may be a linear or nonlinear function.
Classic regression problem studies the relationship between two variables X and Y, where X is called the independent variable and Y is called the dependent variable and we usually register the information about these variables as a pair (X, Y), where X can be a vector or a matrix of values for the independent variables. As a particular case, Simple Linear Model can be adjusted if we assumed that a linear relationship between X and Y is plausible. Under similar assumptions, we consider the calibration problem that is another important interest for researchers. It arises when the focus is on the estimation for a particular value of the independent variable x given an observed value of the dependent variable $Y = y$. Some authors have worked on this topic as (Berkson 1969, Naszódi 1978). They proposed some estimators and discussed their asymptotic properties. Before to present a solution to the calibration function problem from change points, we present a review of the classical and latest calibration problem approaches jointly with the change point problem.

2. Calibration problem

In a particular way, fitting a simple linear regression model (SLRM) implies to quantify the effect of the predictor (X) on the response variable (Y). This is done through the estimation of the model parameters and a posterior residual analysis. After the data set is collected and the model is specified, the next step is to find the estimate the parameters of the model. The general statement for a SLRM is

$$y_i = \alpha + \beta x_i + \varepsilon_i \ , \ i = 1, 2 \ldots, n, \tag{1}$$

with $\varepsilon_i \sim N\left(0, \sigma^2\right) (i.i.d.)$ that corresponds to the usual normality, independence and homoscedasticity assumptions about the random error.
The parameters can be estimated using maximum likelihood or least squares. In both cases the expressions of the parameter estimates are:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \tag{2}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \tag{3}$$

The $\hat{\alpha}$ and $\hat{\beta}$ are Gauss-Markov minimum variance unbiased estimates (Rao 1973, Berkson 1969). From a SLRM we get:

$$X_i = \frac{Y_i - \alpha}{\beta} + \eta_i \ , \ i = 1, 2 \ldots, n, \tag{4}$$

where $\eta_i = \frac{\varepsilon_i}{\beta}$ and $\eta_i \sim N\left(0, \frac{\sigma^2}{\beta^2}\right)$. This model is considered a particular approach to solve the inverse regression problem. However, there are several approaches to this problem used to determinate the best estimator for X. Krutchkoff (1967, 1969), Berkson (1969), Chow & Shao (1990), Halperin (1970), Naszódi (1978) have worked on this topic and they have developed some methods to solve some specific problems associated with the calibration problem in the cross-sectional setting. Many other authors cited by Osborne (1991) have also worked on this topic.
Many disciplines conduct studies in which the primary objective is to make inferences based on a linear or nonlinear relationship between the explanatory and response variables. However, this is not the unique interest of the researchers, an inverse relationship could be important, too. If we are interested in predicting a specific value for X, given a value of Y, then we have a calibration or inverse prediction problem and we need to study some specific concepts and conditions about this. Blankenship et al. (2003) studied some properties about calibration problems.

In the case of longitudinal data,observations are collected over the time and some assumptions about the model given by (1) are wrong and we need to consider some additional covariance structures and take them into account to estimate the parameters and predict t ($\hat{t}$) as accurate as possible.

2.1. Parametric approach

As we stated before, the calibration problem can be deal with in different ways. Some authors have worked on parametric approaches most of them referenced in the Osborne's paper and so many others have been working on this topic as Wu et al. (2001) and Blankenship et al. (2003). A Bayesian approach have been worked by Hoadley (1970), Harville (1974), Hunter & Lamboy (1981), and some other authors. Also Knafl et al. (1984), Carroll et al. (1988), Carlstein (1988), Gruet (1996) and Ding & Karunamuni (2004) worked on a nonparametric approach to this problem. Concordet & Nunez (2000), Schwenke & Milliken (1991), Blankenship et al. (2003) have made some approaches for specific problems in which there are a mixture of linear or nonlinear mixed models and calibration problems in applied sciences. Krutchkoff (1967, 1969) began a wide discussion on the calibration problem under independence assumption in cross-sectional setting. Odén (1973) proposed a methodology to obtain the confidence intervals in reverse regression by proposing to select a function which has some specific properties to guarantee that a fixed proportion of intervals contain the true parameter. Brown (1979) proposed the integrated mean squared error (IMSE) and compared the results with the classical and inverse estimator. Also, he discussed its properties as well, and called this the 'optimal estimator '. Trout & Swallow (1979) contrasted results between classic and inverse regression estimates using confidence bands with both a uniform procedure and Scheffe's procedure. He concluded the uniform procedure is enough efficient and simplier than Scheffe's procedure.

Oman (1984) showed some results about residual analysis in calibration by consider a calibrative distance curve. He proposed this methodology and suggested some of this curve properties. Also, he suggested to make a precise specification about the model to run an adequate residual analysis through some methodologies such as Cook's distance and his proposal. Shukla (1972) returned on the Krutchkoff's papers and discussed some specific details about these results, particularly he wrote about how the number of observations in the design effects the parameter estimates, the mean squared error (MSE) and variance in both cases classical and inverse estimator. Perng & Tong (1974) disscused the results showed by Odén and proposed a sequential procedure for the construction of confidence bands to x. Their work can be taken as an important benchmark to asymptotic results obtained for the calibration problem. Minder & Whitney (1975) showed, using likelihood analysis, the relevance of an informative likelihood to get a more precise estimate to x. Naszódi (1978) discussed how to eliminate the bias in calibration problems while the experimental design is running.

In a similar fashion, Carroll et al. (1988) discussed the results presented in the parametric approach worked by Trout & Swallow and suggested other properties about the nonparametric approach on the work made by Knafl et al. (1984). Chow & Shao (1990) showed the difference between classical and inverse regression through some properties about the relative ratio between the estimates and concluded that the values obtained in each case are not interchangeable. Schwenke & Milliken (1991) made an approach to the calibration problem but in the nonlinear case. Osborne (1991) wrote a paper about the main advances on calibration problems made to date. A modified approach to the calibration problem was made by Dahiya & McKeon (1991) who proposed two additional estimators based on past results to get better estimates than the classic and inverse estimates. Hsing (1999) proved properties about the nearest neighborhood technique applied to inverse regression by using some geometric properties. All their work is a theoretical development in this topic.

Kalotay (1971) proposed a solution of the calibration problem by using an structural model under three assumptions which allow to get, by formal mathematical computation, the structural distribution for the model parameters even if the error does not come from a normal distribution. This methodology is so attractive because it supplies the marginal structural distribution for x and avoids some distributional assumptions. His work was based on a similar framework as the Creasy-Feller theorem. Scheffé et al. (1973) made a wide exposition about the statistical theory of calibration. His work presented some results to the calibration intervals. He proposed a graphical method to get the calibration line and analyzed some properties about the estimates under different assumptions. On the other hand, Brown & Sundberg (1989) worked on the calibration problem but extended it to the multivariate case and proposed a change regression. Their development also considered the time series case and supervised and unsupervised learning processes based on comparative information analysis.

Srivastava & Singh (1989) made an interesting study which presented some properties of the classical and inverse estimator. Their conclusions are based on small disturbance asymptotic (SDA) theory. They suggested that the classical estimator is better because it is consistent. They studied some additional properties about the classic and the inverse estimator, by analyzing the Asymptotic Mean Squared Error (AMSE) and variance. However, their study was limitated to small values of the variance of the errors, that is, the random errors in the calibration are relatively small. Schwenke & Milliken (1991) explored the properties about the classical estimator for the calibration problem but in the nonlinear model case by constructing the confidence bands using the distribution on the $\hat{x}$ and the distribution on the $\hat{\beta}$. This work also presented a methodology to test the equality of two calibration points by using different monotone functions, f and g,respectively over the regression range. Based on simulation studies, they showed that the asymptotic testing procedure performed well small sample sizes.

Kimura (1992) compared the estimates of an unknown value of x by considering two models and in conjunction with the Expectation Maximization (EM) algorithm to get estimates under each model and constrast them against the eigenvector estimators including a large sample approximation to obtain the estimates for standard errors. He suggested that the Eigenvectors Estimators (EV) and the EM methods had an identical maximum likelihood estimate of regression coefficients. Denham & Brown (1993) worked on calibration with serially correlated errors. They considered the estimation of a matrix $\Gamma = \sigma^2\mathbf{I}$ to allow more flexibility on the covariance matrix. They also considered a stationary autocorrelation process and adapted the Box-Jenkins procedure to avoid some difficulties to obtain the generalized least squares estimator and worked on the calibration problem using cubic splines. They illustrated the methodology by using a data set on analysis of the infrared absorption in detergents and remarked that 'methods of estimation based on many frequencies retain better robustness in routine use under varying experimental conditions '. Their work has a plus because of it was the first approach to calibration using linear mixed models.

Extending the works made before by some authors, Wang et al. (1997) proposed an estimator and obtained its asymptotic properties in presence of censored data. They performed a simulation study allowing a fixed percentage of censored data and checked the change in bias, MSE, and standard error of the estimates. They found these methods much more efficient than similar methods presented in the available literature.

A robust calibration method was presented by Cheng & Van Ness (1997). They compared the estimators obtained using several regression methods. They showed that robust calibration methods have a good behavior in presence of outliers and with non-normal errors distribution. Also, they suggested an outlier detection procedure to identify simple-outliers in calibration problems.

Sundberg (1999) showed an application to the multivariate calibration problem by studying both univariate and multivariate estimation and he provided a complete review on the estimators in linear and non-linear case. He suggested some confidence regions and diagnostics in the calibration problem and illustrated them by using some data from a chemical study. Furthermore, he suggested to evaluate the goodness of fit by comparing the estimators obtained under several models, including ridge regression.

2.2. Nonparametric approach

Knafl et al. (1984) made a nonparametric approach to the calibration problem and obtained some confidence bands under nonmonotonic functions. However, they suggested the effectiveness of their procedures depends on the accuracy of the estimates. They also mentioned that calibration problems are strongly related to regression problems. They provided an useful algorithm to estimate a nonparametric calibration function and its confidence bands. They illustrated this methodology and gave some remarks about this procedure.

Benton et al. (2003) proposed a parametric bootstrap method to test the hypothesis $H_0 : x \leq c$ against $H_a : x > c$, where c is a known value given to y_0 under the selected model. They proposed to use a modified pivot statistic to get confidence regions based on it. They compared the power of two tests based on this statistic for both univariate and multivariate one-sided hypothesis testing to the calibration problem and made some additional considerations. They suggested to use a parametric bootstrapping method to improve the quality of the inferences on the calibration problem.

2.3. Bayesian approach

Racine-Poon (1988) presented a Bayesian approach to the calibration problem. He proposed to obtain the estimates of the parameters taking into account the prior information given by the calibration experiment. Also, they assumed that the random errors come from a normal distribution. He obtained the posterior distribution under the parameters independence condition and by consider a particular value $\eta = x$ to estimate. He considered a logit model to describe the relationship between concentration and the response variable. He suggested that the posterior distributions have a normal distribution in many calibration experiments. He also suggested to avoid an improper posterior distribution by considering only proper a prior distributions. DeJong et al. (1996) presented an alternative approach to deal with the calibration problem. They proposed a symmetric treatment for researcher's uncertainty according to the model and an additional empirical model by considering prior distributions for both theoretical and empirical models. They also dealt with serial correlation but these results are not appealing neither of the prior distributions considered by them. They applied the methodology using technology market data and showed the weakness and strenght of the model and its sensitivity to the specification of the parameter uncertainty. Gruet (1996) proposed an additional method to estimate a quantity of interest by combining kernel (classic nonparametric) and robust estimation techniques. She developed a theory and showed some main results and properties as consistency of the estimator proposed and some other important characteristics in multivariate calibration. She also considered simultaneous tolerance sets to build simultaneous calibration intervals.

On the other hand, Lucy et al. (2002) presented a nonparametric approach to the calibration problem. They used a Bayesian and an empirical Bayesian estimation procedure in the calibration problem. Also, they worked on a nonparametric approach using smoothing on continuous and discrete variables and contrasted the obtained values using some measurements as the Mean Absolute Deviation (MAD), systematic bias by estimating the slope coefficient against the known x value and the mean width of the 95 % confidence intervals. These approaches were illustrated using on forensic data. They found that the smoothed empirical Bayesian calibration method yielded estimates with accuracy and precision similar to the multiple regression but it is better due to its robustness against the systematic bias, particularly to extreme values in the distribution of X. Ding & Karunamuni (2004) proposed a new estimator for the calibration problem. They compared classical, inverse and their estimator using the MSE. They also showed some additional asymptotic properties of this estimator under some scenarios. They performed a simulation study which considered X as a random variable and suggested that the proposed estimator is as good as the inverse or the classic estimator under some conditions.

2.4. Extension to linear and nonlinear mixed models

Developments on the calibration problems have been widely made in the last decades. Næs (1985) made an approach to the calibration problem but in the multivariate case. He proposed an adapted estimator to obtain a prediction of X under multivariate normality assumptions and also worked on multivariate calibration on the specific case in which the covariance matrix has a predetermined structure. He proposed a multivariate calibration method finding that the new predictor coincides with the Multiple linear regression predictor for X on Y when the factors and the amount of variables are similar. He pointed out that the main idea of this approach is to estimate the covariance matrix for Y in a better way than usual. This work is important because it has an specific development when the collected data presents multicollinearity problems. Furthermore, some extensions considered by some authors included the calibration problem under a longitudinal setting approach and using linear mixed models.

Fornell et al. (1991) investigated some properties of the direct and reverse regression when some unobservable variables are including in the model. They concluded that the reverse regression had some limitations compared to direct regression, because the estimates obtained by using it are generally biased. Also, they proposed an alternative method to obtain unbiased estimates. This method is based on linear mixed models which consider a flexible covariance structure for the errors.

3. Change point problem

Fitting a simple linear regression model to a data set in a cross-sectional setting is a common practice. It is usually assumed that the considered model holds for the whole data. However, sometimes researchers need to consider linear models where the structure of the model changes. An exploratory data analysis could allow to detect a change in the model structure in either any specific point or several points. The point in which the structure changes is called the change point. If we consider a model with a single continuous change point, the model given by (1) can be rewriten as:

$$y_i = \begin{cases} \beta_{10} + \beta_{11}x_i + \varepsilon_{1i} & i = 1, \ldots, s \qquad \varepsilon_{1i} \sim N\left(0, \sigma_1^2\right) \\ \beta_{20} + \beta_{21}x_i + \varepsilon_{2i} & i = s+1, \ldots n \qquad \varepsilon_{2i} \sim N\left(0, \sigma_2^2\right) \end{cases} . \tag{5}$$

In general, change points can be known or unknown and they divide a statistical model into homogeneous segments. Usually, the statistical model can be fitted as a piecewise or broken-stick regression. The former is applied when the model has a discontinuity at a specific point (change point) and the latter when the model has a continuous change point. In statistical inference a change point exists if there is enough evidence against the null hypothesis of 'no change'. Several authors worked on this problem, Carlstein (1988), Muller (1992), Hartigan (1994), Bhattacharya (1994), Küchenhoff (1996), Jandhyala & MacNeill (1997), Neumann (1997), Rogers (2010), Saatçi et al. (2010), Killick & Eckley (2014),Chen & Gupta (2000) and some others. These authors worked on this problem in cross-sectional setting under independence assumption.

Hofrichter (2007), based on McCullagh & Nelder (1989) works on a parametric approach about the change point estimation by considering Generalized Linear Models (GLM). The results from the simulation study showed that and the results showed that this approach was so effective to detect one single change point or multiple change points. He built an R-package called *CpInGLM* but unfortunately, it is not still available at the CRAN mirror. Zhou et al. (2008) also works under this setting but in their approach they suggested to modify the objective function to eliminate the non-smothness problem with the change point problem in the maximum likelihood function.

Some algorithms useful for the detection of change points are available. detection can be structured. We have explored some papers and working with an R-project package called *changepoint*. This package was proposed by Killick & Eckley (2014) and this methodology is particularly useful when time series are the object of study. They worked on large samples and considered changes on the mean and the variance which allow a better estimation. The detection of a single change point can be presented as a hypothesis test. The null hypothesis, H_0, corresponds to no change point ($m = 0$) and the alternative hypothesis, H_1, is a single change point ($m = 1$) (Chen & Gupta 2000). However this idea is not new, some advances were done in this topic, for instance, Farley & Hinich (1970) developed a test of the null hypothesis that a slope coefficient in a time series model does not shift, against the alternative that the parameter shifts exactly once and the potential shift is small relative to the error variance.

Bai & Perron (2003) introduced a dynamic programming algorithm to identify how many change points had a model under a linear structure and to estimate these change values. They suggested to build first a triangular matrix of sums of squared residuals and then sequentially evaluate the each partition of data until the new partition achieves a global minimization on the overall sum squared regression. They also suggested a modified algorithm to evaluate a partial structural change and mentioned some properties of this algorithms including the convergence rates. They worked on confidence regions for these change points and the parameters estimates. They also performed a study to determine the number of partitions that should be considered in each model by consider a test of no break against a fixed number of breaks. Downey (2008) proposed a Bayesian algorithm to find the change points which they called an online technique that allows to identify change points when the evidence of this reaches a certain threshold. He proposed to evaluate this technique using a generalized linear model procedure based on the cummulative distribution function and a nonparametric procedure that allows to calculate the change point probability, and finally he found some specific properties associated to each one.

3.1. Nonparametric approach to the change point problem

Carlstein (1988) made a nonparametric aproach to the change point problem and he found a set of consistent estimators. He studied the rates of convergence and the error probability associated to each one. He also remarked that the error probability has an exponencial bound and proved it through four lemmas which were exposed and proved on his paper. Simmilarly, Muller (1992) suggested some change points estimators and established the conditions under which them would be consistently estimated. His work was made through a comparison of left and right one-sided kernel smoothers and also included the detection of discontinuities. He showed that the rates of convergence are as similar as when the change point is known. Hartigan (1994) studied the change point problem by evaluate the performance of linear estimators under the assumption that the parameter values did not vary smoothingly. In this case, the linear estimators could be wrong because if the parameter of interest is so near to a discontinuity then the weighted sum will include an specific kind of bias. This bias was introduced by the observations on the other side of the discontinuity and modifies the value of the parameter estimators. He considered a lower bounds for the minimax risk in three cases. First one, by consider a circular change point problem on an even number of data. The second one, aditionally considered the linear estimator as the shift estimator which result useful for adaptative methods but it is not in practice. At the last one, he considered linear estimators for image segmentation problems.

Darkhovski (1994) worked the change point problem using nonparametric methods and he studied the problem by consider two ideas, the first idea is based on consider the problem of detection of changes in the mean value of some new sequences and the second idea was based on Kolmogorov-Smirnov statistics that allows to detect the change point in these sequences. He presented both methods and some concrete algorithms to find the change point and exposed some conditions to guarantee the almost sure convergence of these sequences. Simmilarly, Hušková & Picek (2005) made an approach to this problem by a bootstrap technique and they explained the main results of their proposal through theoretical results. They proved the theorems applied on the change point problem under asymptotical assumptions. On the other hand, change point problem is not only a regression problem, stocastic processes is an important field where change problem is important, too. Dayanik et al. (2008) worked a change detection under a Bayesian approach but considering a Wiener and a Poisson process and solved optimal stopping problems for jump-diffusion processes by separating jump and diffusion parts with the help of a jump operator. Their work could be seen as a solution to a particular problem on manufacturing processes or production processes because in both cases we need to identify promtly and accurately the changes to avoid that the process goes out of control or false alarm signals.

4. Change points and linear mixed models

Linear Mixed Models (LMMs) have been widely studied and they are usually seen as an extension of the model (1), but in this case taking into account that the random errors can be divided into two parts: the former, between-subjects variability and the latter within-subjects variability. The general expression for LMMs (Fitzmaurice et al. 2008), is given by:

$$\begin{aligned} \mathbf{Y}_i = \quad & \underbrace{\mathbf{X}_i\boldsymbol{\beta}}_{\text{Fixed}} + \underbrace{\mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i}_{\text{Random}} \\ & \mathbf{b}_i \sim N\left(\mathbf{0}, \mathbf{D}\right) \\ & \boldsymbol{\varepsilon}_i \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_i\right) \\ & \boldsymbol{\varepsilon}_i \text{ and } \mathbf{b}_i \text{ are independent.} \end{aligned} \tag{6}$$

Most recent works on change points in linear mixed model had been limited only to some authors as Lai & Albert (2014) and Rosenfield et al. (2010) which considered change points as a unique point through the blocks or factors. They rewrote the model (6)as:

$$\mathbf{Y}_{ij} = \sum_{k=1}^{g+1} \mathbf{X}_{ij} \boldsymbol{\beta}_{B_k} I\left(t_{ij} \in B_k\right) + \mathbf{Z}_{ij} \mathbf{b}_i + \mathbf{e}_{ij} \tag{7}$$

and in case of a fixed effects and random effects model it can be written as:

$$\mathbf{Y}_{ij} = \sum_{k=1}^{g+1} \mathbf{X}_{ij} \boldsymbol{\beta}_{B_k} I\left(t_{ij} \in B_k\right) + b_{i,0} + \sum_{k=1}^{g+1} b_{i,B_k} I\left(t_{ij} \in B_k\right) + \mathbf{e}_{ij}, \tag{8}$$

where g corresponds to the number of blocks. They worked the problem as a model with a common change point by building blocks until a particular value of the fixed variable. However, they did not predict an individual change point for each subject under a longitudinal setting. They neither worked on the prediction of an specific time given some previous conditions on the fixed variable. Predicting a time in which the model changes is so important in productive processes because, for example, this could allow to avoid some additional drawbacks, especifically it could help to reduce the storage expenses. Jackson & Sharples (2004) worked on the change point problem for longitudinal data considering these as censored change points. They worked on a Bayesian approach to this problem and studied the change point problem by consider two models, the first one was defined as a linear progression and it is related to those subjects who showed a smooth profile on its measures through the time. The second one was defined as the acute onset model that allows to explain a sudden change on the subject-specific profile. Their work was illustrated using a particular data set about lung transplantation.

5. Conclusions

We have reviewed most of the literature about this topic and we have found an extensive work and some useful methodologies. The proposal of this review was to establish the main references and advances that has been made on this topic to ensure the relevance and the appropriateness of our research proposal (Garcia et al. 2015) which consider as the main goal to estimate the unknown change points for each subject in longitudinal studies by using a linear mixed models approach and once we know these change points we want to build a calibration function that allows us to predict the change point according with the available information about the fixed effects considered at the onset modeling stage. We will propose a methodology useful in that research fields which it is important to predict a specific point. For example, predict the time when a person becomes healthier according to the specific conditions of a treatment, or establish the maximum time that a wooden slot should be dryed before selling it and thus not increasing the storage expenses.

We consider Hofrichter (2007)'s work as a benchmark to generalize the change point identification under a longitudinal setting and the Lai & Albert (2014)'s work as one of the most important references to estimate the change points for Linear Mixed Models. The main results of that research will be presented later.

Recibido: 11 de junio de 2016
Aceptado: 21 de noviembre de 2016

References

Bai, J. & Perron, P. (2003), 'Computation and analysis of multiple structural change models', *Journal of applied econometrics* **18**(1), 1–22.

Benton, D., Krishnamoorthy, K. & Mathew, T. (2003), 'Inferences in multivariate–univariate calibration problems', *Journal of the Royal Statistical Society: Series D (The Statistician)* **52**(1), 15–39.

Berkson, J. (1969), 'Estimation of a linear function for a calibration line; consideration of a recent proposal', *Technometrics* pp. 649–660.

Bhattacharya, P. (1994), 'Some aspects of change-point analysis', *Lecture Notes-Monograph Series* **23**, 28–56.

Blankenship, E. E., Stroup, W. W., Evans, S. P. & Knezevic, S. Z. (2003), 'Statistical inference for calibration points in nonlinear mixed effects models', *Journal of Agricultural, Biological, and Environmental Statistics* **8**(4), 455–468.

Brown, G. (1979), 'An optimization criterion for linear inverse estimation', *Technometrics* **21**(4), 575–579.

Brown, P. J. & Sundberg, R. (1989), 'Prediction diagnostics and updating in multivariate calibration', *Biometrika* **76**(2), 349–361.

Carlstein, E. (1988), 'Nonparametric change-point estimation', *The Annals of Statistics* **16**(1), 188–197.

Carroll, R., Spiegelman, C. & Sacks, J. (1988), 'A quick and easy multiple-use calibration-curve procedure', *Technometrics* **30**(2), 137–141.

Chen, J. & Gupta, A. K. (2000), *Parametric Statistical Change Point Analysis*, Birkhauser.

Cheng, C.-L. & Van Ness, J. W. (1997), 'Robust calibration', *Technometrics* **39**(4), 401–411.

Chow, S.-C. & Shao, J. (1990), 'On the difference between the classical and inverse methods of calibration', *Applied statistics* **39**(2), 219–228.

Concordet, D. & Nunez, O. G. (2000), 'Calibration for nonlinear mixed effects models: an application to the withdrawal time prediction', *Biometrics* **56**(4), 1040–1046.

Dahiya, R. C. & McKeon, J. J. (1991), 'Modified classical and inverse regression estimators in calibration', *Sankhyā: The Indian Journal of Statistics, Series B* **53**(1), 48–55.

Darkhovski, B. S. (1994), 'Nonparametric methods in change-point problems: A general approach and some concrete algorithms', *Lecture Notes-Monograph Series* pp. 99–107.

Dayanik, S., Poor, H. V. & Sezer, S. O. (2008), 'Multisource bayesian sequential change detection', *The Annals of Applied Probability* pp. 552–590.

DeJong, D. N., Ingram, B. F. & Whiteman, C. H. (1996), 'A bayesian approach to calibration', *Journal of Business & Economic Statistics* **14**(1), 1–9.

Denham, M. & Brown, P. (1993), 'Calibration with many variables', *Applied Statistics* pp. 515–528.

Ding, K. & Karunamuni, R. J. (2004), 'A linear empirical bayes solution for the calibration problem', *Journal of statistical planning and inference* **119**(2), 421–447.

Downey, A. B. (2008), 'A novel changepoint detection algorithm', *arXiv preprint arXiv:0812.1237* .

Farley, J. U. & Hinich, M. J. (1970), 'A test for a shifting slope coefficient in a linear model', *Journal of the American Statistical Association* **65**(331), 1320–1329.

Fitzmaurice, G. M., Laird, N. M. & Ware, J. H. (2008), *Applied longitudinal analysis*, Vol. 998, John Wiley & Sons.

Fornell, C., Rhee, B.-D. & Yi, Y. (1991), 'Direct regression, reverse regression, and covariance structure analysis', *Marketing Letters* **2**(3), 309–320.

Garcia, E. K., Correa, J. C. & Salazar, J. C. (2015), 'Estimación de una función de calibración completa a partir de puntos de cambio', *Comunicación corta, Memorias XXV Simposio Internacional de Estadística 2015* .

Gruet, M.-A. (1996), 'A nonparametric calibration analysis', *The Annals of Statistics* **24**(4), 1474–1492.

Halperin, M. (1970), 'On inverse estimation in linear regression', *Technometrics* **12**(4), 727–736.

Hartigan, J. (1994), 'Linear estimators in change point problems', *The Annals of Statistics* **23**, 824–834.

Harville, D. A. (1974), 'Bayesian inference for variance components using only error contrasts', *Biometrika* **61**(2), 383–385.

Hoadley, B. (1970), 'A bayesian look at inverse linear regression', *Journal of the American Statistical Association* **65**(329), 356–369.

Hofrichter, J. (2007), *Change point detection in generalized linear models*, Graz University of Technology. doi:http://www.stat.tugraz.at/dthesis/Hofrichter07.pdf. Accessed on: 2015/04/22.

Hsing, T. (1999), 'Nearest neighbor inverse regression', *Annals of statistics* **27**(2), 697–731.

Hunter, W. G. & Lamboy, W. F. (1981), 'A bayesian analysis of the linear calibration problem', *Technometrics* **23**(4), 323–328.

Hušková, M. & Picek, J. (2005), 'Bootstrap in detection of changes in linear regression', *Sankhyā: The Indian Journal of Statistics* pp. 200–226.

Jackson, C. H. & Sharples, L. D. (2004), 'Models for longitudinal data with censored changepoints', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **53**(1), 149–162.

Jandhyala, V. & MacNeill, I. (1997), 'Iterated partial sum sequences of regression residuals and tests for changepoints with continuity constraints', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **59**(1), 147–156.

Kalotay, A. (1971), 'Structural solution to the linear calibration problem', *Technometrics* **13**(4), 761–769.

Killick, R. & Eckley, I. (2014), 'changepoint: An r package for changepoint analysis', *Journal of Statistical Software* **58**(3), 1–19.

Kimura, D. K. (1992), 'Functional comparative calibration using an em algorithm', *Biometrics* pp. 1263–1271.

Knafl, G., Spiegelman, C., Sacks, J. & Ylvisaker, D. (1984), 'Nonparametric calibration', *Technometrics* **26**(3), 233–241.

Krutchkoff, R. (1967), 'Classical and inverse regression methods of calibration', *Technometrics* **9**(3), 425–439.

Krutchkoff, R. (1969), 'Classical and inverse regression methods of calibration in extrapolation', *Technometrics* **11**(3), 605–608.

Küchenhoff, H. (1996), 'An exact algorithm for estimating breakpoints in segmented generalized linear models'.

Lai, Y. & Albert, P. (2014), 'Identifying multiple change points in a linear mixed effects model', *Statist. Med* **33**, 1015 – 1028. doi: 10.1002/sim.5996.

Lucy, D., Aykroyd, R. & Pollard, A. (2002), 'Nonparametric calibration for age estimation', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **51**(2), 183–196.

McCullagh, P. & Nelder, J. A. (1989), *Generalized linear models*, Vol. 2, Chapman and Hall London.

Minder, C. E. & Whitney, J. (1975), 'A likelihood analysis of the linear calibration problem', *Technometrics* **17**(4), 463–471.

Muller, H.-G. (1992), 'Change-points in nonparametric regression analysis', *The Annals of Statistics* **20**(2), 737–761.

Næs, T. (1985), 'Calibration when the error covariance matrix is structured', *Technometrics* **27**(3), 301–311.

Naszódi, L. J. (1978), 'Elimination of the bias in the course of calibration', *Technometrics* **20**(2), 201–205.

Neumann, M. H. (1997), 'Optimal change-point estimation in inverse problems', *Scandinavian Journal of Statistics* **24**(4), 503–521.

Odén, A. (1973), 'Simultaneous confidence intervals in inverse linear regression', *Biometrika* **60**(2), 339–343.

Oman, S. D. (1984), 'Analyzing residuals in calibration problems', *Technometrics* **26**(4), 347–353.

Osborne, C. (1991), 'Statistical calibration: a review', *International Statistical Review/Revue Internationale de Statistique* **59**(3), 309–336.

Perng, S. & Tong, Y. L. (1974), 'A sequential solution to the inverse linear regression problem', *The Annals of Statistics* **2**(3), 535–539.

Racine-Poon, A. (1988), 'A bayesian approach to nonlinear calibration problems', *Journal of the American Statistical Association* **83**(403), 650–656.

Rao, C. R. (1973), *Linear statistical inference and its applications*, Vol. 2nd.Ed., John Wiley & Sons.

Rogers, A. N. (2010), 'A nonparametric method for ascertaining change points in regression regimes'.

Rosenfield, D., Zhou, E., Wilhelm, F. H., Conrad, A., Roth, W. T. & Meuret, A. E. (2010), 'Change point analysis for longitudinal physiological data: detection of cardio-respiratory changes preceding panic attacks', *Biological psychology* **84**(1), 112–120.

Saatçi, Y., Turner, R. D. & Rasmussen, C. E. (2010), Gaussian process change point models, *in* 'Proceedings of the 27th International Conference on Machine Learning (ICML-10)', pp. 927–934.

Scheffé, H. et al. (1973), 'A statistical theory of calibration', *The Annals of Statistics* **1**(1), 1–37.

Schwenke, J. R. & Milliken, G. A. (1991), 'On the calibration problem extended to nonlinear models', *Biometrics* **47**(2), 563–574.

Shukla, G. (1972), 'On the problem of calibration', *Technometrics* **14**(3), 547–553.

Srivastava, V. & Singh, N. (1989), 'Small-disturbance asymptotic theory for linear-calibration estimators', *Technometrics* **31**(3), 373–378.

Sundberg, R. (1999), 'Multivariate calibration: direct and indirect regression methodology', *Scandinavian Journal of Statistics* **26**(2), 161–207.

Trout, J. R. & Swallow, W. H. (1979), 'Regular and inverse interval estimation of individual observations using uniform confidence bands', *Technometrics* **21**(4), 567–574.

Wang, C., Hsu, L., Feng, Z. & Prentice, R. L. (1997), 'Regression calibration in failure time regression', *Biometrics* pp. 131–145.

Wu, W. B., Woodroofe, M. & Mentz, G. (2001), 'Isotonic regression: Another look at the changepoint problem', *Biometrika* **88**(3), 793–804.

Zhou, H., Liang, K.-Y. et al. (2008), On estimating the change point in generalized linear models, *in* 'Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen', Institute of Mathematical Statistics, pp. 305–320.