Incorporation of environmental covariates to nonlinear mixed models describing fruit growth

December 2023

Del Brio, D.[1]; Tassile, V.[2]; Bramardi, S.J.[3]; Reeb, P.D.[3]

ABSTRACT

Yield prediction is still a major challenge in pear production. Forecasting fruit growth after modeled curves allows predicting both potential yield and quality. This research aimed to fit multilevel no-linear mixed models (NLMM) based on logistic curves to describe pear growth in the Upper valley of Rio Negro and Neuquén, Argentina. The models incorporated several thermo-accumulative indices accounting for temperature effects on fruit-growth physiology. In this way, they captured normal fruit-growth patterns and environmental variations along the growing season. The study was conducted on "William´s" pear trees for 16 seasons. Many trees and fruits were randomly selected and identified. Equatorial diameters were weekly measured with an electronic digital caliper. Climatic data was recorded for all studied seasons from INTA Upper Valley agrochemical station and thermo accumulative indexes were calculated from daily temperature. The best models were selected according to the information criteria index. Multilevel NLMM discerned and quantified the sources of stochastic variability at different levels, allowing better index criteria in comparison to models only considering a single level of variability among random effects. The incorporation of thermo accumulative indexes also increased model performance.

Keywords: pear, growth curves, yield prediction, random effects.

RESUMEN

El pronóstico de cosecha es un gran desafío en la producción de peras. Estimar el tamaño de los frutos a partir de curvas de crecimiento permite predecir tanto la cantidad como la calidad de la fruta para cosecha. Este trabajo tuvo como objetivo ajustar modelos mixtos no lineales multiniveles (MMNL) basados en la curva logística para describir el crecimiento de peras "William´s" en el Alto Valle de Río Negro y Neuquén, Argentina. Los modelos ajustados incorporaron diferentes índices termoacumulativos que contemplan los efectos de la temperatura en la fisiología del crecimiento de los frutos. De esta manera, se logra no solo describir el crecimiento de los frutos, sino también se pueden observar las variaciones ambientales a lo largo de las temporadas de crecimiento. El estudio se realizó en perales "William´s" durante 16 temporadas. Se seleccionaron e identificaron al azar numerosos árboles y frutos. A cada fruto se le midió su diámetro ecuatorial semanalmente con un calibre digital electrónico. Los datos climáticos se obtuvieron de la estación meteorológica del INTA Alto Valle y los índices termoacumulativos se calcularon a partir de los datos de temperaturas. Los mejores modelos fueron seleccionados según los criterios de información. El MMNL multinivel permitió discernir y cuantificar las fuentes de variabilidad estocástica en diferentes niveles, lo que permitió obtener mejores criterios de información en comparación con los modelos que solo consideraron un único nivel de variabilidad entre los efectos aleatorios. La incorporación de índices termoacumulativos mejoró notablemente la performance de los modelos obtenidos.

Palabras clave*: peras, curvas de crecimiento, pronóstico de producción, efectos aleatorios.*

[1]Instituto Nacional de Tecnología Agropecuaria (INTA), Estación Experimental Agropecuaria (EEA) Alto Valle, Ruta Nacional N 22, km 1190, (8332), Allen, Río Negro, Argentina. Correo electrónico: delbrio.dolores@inta.gob.ar
[2]Universidad Nacional del Comahue, Facultad de Ciencia y Tecnología de los Alimentos, 25 de Mayo y Reconquista, (8336), Villa Regina, Río Negro, Argentina.
[3]Universidad Nacional del Comahue, Departamento de Estadística, Buenos Aires 1400, (8300) Neuquén, Argentina.

Doi https://doi.org/10.58149/14h1-sp68

INTRODUCTION

Yield prediction is considered vital information for growers as it defines management and marketing strategies. In many commercial orchards yield prediction has become a frequent practice, including total production and fruit size distribution. This information enables farmers to improve packaging, storage and marketing decisions. In marketing, differences of a few millimeters in fruit size have great implications on money returns (Giménez and Tassile, 2015).Until now, there were two main methods for predicting fruit size at harvest: stochastic models and curves. Stochastic models are based on the assumption of a relation between tree fruit-load and final size. Growth curves attempt to describe fruit growth patterns through nonlinear, particularly sigmoid models. These curves provide an upper asymptote indicating maximum reachable sizes, an inflection point showing maximum variation rate and a lower asymptote providing initial fruit size (Alvarez and Boche, 1999).

Different functions have been proposed to describe apple growth. A double sigmoid curve considering three well-defined phases (Magein, 1989), an expolinear model (Lakso *et al*., 1995), a logistic model (Ortega Farias *et al*. 1998), and a negative exponential function (Zadravec *et al*., 2014). As regards pears, fruit growth has been well described by logistic curves in "Williams" cultivar (Bramardi *et al*., 1997; Giménez, 2012) in which the third parameterization has been proved the most appropriate (Ratkowsky, 1983).

Models fitting growth patterns of a specific cultivar in a certain region are not generalizable to other crop conditions. Final fruit size depends on many factors, such as genotype, crop management, weather, and fruit number (Bramardi *et al*., 2006; Jiménez and Royo Díaz, 2003). Climate directly and indirectly affects growth, yield and quality of crops (Oliveiora Aparecido *et al*., 2016).

Several studies have shown the relationship between weather and yield in different crops. For instance, maximum and minimum temperatures affect grain crop yields (Kumar *et al*., 2014); temperature, precipitation and pollen emission are used to estimate olive production (Galan *et al*., 2004); meteorological variables are used to predict fruit number per tree in oranges (Pasqua *et al*., 2007); and degree-days and potential evaporation influence apple diameter (Kaack and Pedersen, 2010).

Incorporating environmental variables into growth curves becomes possible through non-linear mixed models (NLMM), characterizing typical fruit-growth patterns, and contemplating the possibility of incorporating random effects according to clustering levels, modelling the correlation error, and integrating covariates that characterize the growing season. In addition, NLMM are easier to interpret, consider parsimony, and validate regardless of the range of observed data (Pinheiro and Bates, 2000).

Fixed effects of the structural components in NLMM are based on longitudinal data obtained from equatorial diameters of certain fruits throughout their growth cycle. However, in this way, random effects estimation is conditioned to restrictions imposed by the initial fruit selection. If the models have predictive purposes, it is necessary to generalize those to a larger population. This is achieved by considering transversal measurements defined as random samples of equatorial fruit diameters in a reduced number of growth cycle moments (Tassile *et al*., 2015).

Moreover, error correlation structure needs to be included since randomizing the time factor is not possible for repeated measurements over time on the same fruit.

The inclusion of random parameters on different clustering levels could significantly improve model fit. Furthermore, stochastic residual variability present in a significant magnitude indicates that some fixed effects like genetics, topography, or climate factors have been ignored by the model (Calama and Montero, 2005). Information on these factors is not usually available in growth data, therefore, random effects can approximate these fixed effects, either unknown or difficult to quantify.

Random effects are also useful to obtain Empirical Bayesian Estimators (EBEs). EBEs can examine groups of interest, subpopulations, or reveal the existence of quantifiable fixed effects not considered but detected through associations between these and possible sources of variability in any of the levels considered of the random effects. Environmental covariables arise here as good alternatives to characterize the effects of site and growth cycle.

The literature cites several studies based on NLMM explaining pear growth considering correlated data and random effects (Tassile *et al*., 2002), blueberry growth analyzing three different cultivars in low temperature environments (Godoy *et al*., 2008); prediction of crop yield of young eucalyptus plantations and description ofsite quality (Carrero *et al*., 2008), among others. This study aimed to include random effects and environmental covariates to NLMM for example, explaining pear growthin "Williams" pears in the Upper Valley of Río Negro and Neuquén.

MATERIALS AND METHODS

Longitudinal and Transversal data

This study was conducted in seven "Williams" pear orchards located in the upper valley of Río Negro and Neuquén, during 16 seasons for longitudinal data (9 seasons for the modeling base and 7 seasons for the validation base). The orchards are managed under usual regional practices in terms of pruning, sanitary treatments, thinning, and fertilization. Therefore, the models were fit to trees with optimal fruit load preventing fruit load from being considered a factor affecting them.

Considering each orchard and season, some trees were randomly selected and fruits of three size strata (small, medium and large) were carefully chosen. To define size strata, maximum and minimum size of 250 fruits was recorded and size range was calculated and later divided into small, medium and large categories. In total, considering all seasons, 46 trees and 1,599 individualized fruits were measured, from which 17,667 records of equatorial diameter were obtained.

Each fruit was identified and its equatorial diameter was weekly measured with an electronic digital caliper (Essex Stainless Hardened). Measurement moments are referenced as days after full bloom (DAFB), starting 25-30 DAFB with average diameters of 10-15 mm and finishing at harvest time. Two different sides of the equatorial zone of the fruit were considered. Each fallen fruit was replaced by another of similar size and the same tree.

Transversal data were obtained from eight sites that arise from the possible combinations between four seasons and two orchards. In December, at approximately 70 DAFB, some pear trees were selected and equatorial diameter of all their fruits was measured. Then, at harvest, those fruits were collected recording equatorial diameter and weight.

Environmental data

In order to consider potential environmental covariates, daily average temperature was calculated from maximum and minimum temperature recorded from August 1st to January 31st for each of the 16 seasons considered. Data was obtained at the agroclimatic station of INTA Alto Valle.

From daily temperature, 976 thermal accumulation indices (ITAC) were generated for different moments in the season, and 594 ITAC for December in all seasons. These indixes depend on varying criteria and according to accumulation starting point, accumulation completion, and accumulation base temperature.

Data analysis

Model fittings were done using NLME and LMER library of R software (R CORE TEAM; 2018). Likelihood is obtained by Taylor expansion according to the Best Linear Unbiased Estimates (BLUPs) linearization method proposed by Lindstrom and Bates (1990) to obtain NLME estimates with random factors. Models obtained by NLME present nested effects, or a very simple cross-effects structure, allowing direct heteroscedasticity modelling and residual correlation.

Then, the best model was selected based on the Information Criteria Index (ICI). The significance of fixed effects was evaluated by t-tests, based on Wald statistics with degrees of freedom obtained from the Between-Within method.

The NLMER function of the R package lme4 version 1.1-7 was used to obtain estimations. This function calculates likelihood from LAPLACE approximation for multilevel NLME, both nested and crossed.

When necessary, dummy variables were generated to model the fixed effects of the parameters, when it was impossible to obtain them directly from the NLMER function.

Once the estimation process was finished, the variance component structure was evaluated by the ICI, while fixed effects inference was made from likelihood ratio tests (LRT).

Since the obtained models have predictive objectives, the final choice is conditioned by the performance of the predictive criteria (PC) obtained for each candidate model.

The general environmental covariate base was built with a function obtaining representative values for each season based on TCIs. Environmental covariate candidates were selected from the correlations between random effects EBEs at site level and the general environmental covariate. The final selection consists of visual insights into the distribution of candidate environmental covariates and their relation to the EBEs under consideration.

Selected environmental covariates are considered in the construction of new NLME as fixed effects and considering present and/or absent terms in the random effects variance structure.

RESULTS AND DISCUSSION

Fitted model from longitudinal data

The obtained models forecast fruit size as equatorial diameter according to the DAFB, unlike other authors who suggest measuring the fruit weights or volumes (Lakso *et al*., 1995). In particular, growth is described by a logistic curve in third parameterization, agreeing with Bramardi *et al*. (1997) and Giménez (2012) who studied pears cv. 'Williams' and 'Packham's Triumph' in the same region. Other authors have adjusted double sigmoid curves (Magein, 1989), expolinear models (Lakso *et al*., 1995), negative exponential functions (Zadravec *et al*., 2014), or quadratic functions (Atay *et al*., 2010).

The factors used to fit the logistic model were SITE (as the combination between seasons and orchards); TREES (which were randomly selected in each SITE); FRUIT (which were randomly selected for each TREE from 3 strata of SIZES). Based on this, the proposed model (equation 1) included SIZE as fixed effect for parameters of the structural components and random effect generated for FRUIT, TREE, SIZE*TREE and SITE. The least was incorporated to extend the inference space beyond the observed range.

Equation 1

$$Y_{Jklmn} = \frac{1}{(\beta_{1,\,SIZE_k} + b_{1,jklm}) * 0.01 + ((\beta_{2,\,SIZE_k} + b_{2,jklm}) * 0.01) * ((\beta_{3,\,SIZE_k} + b_{3,jklm}) * 0.1)^{DAFB}} + \varepsilon_{jklmn}$$

With

$$b_{i\,,jklm} = b_{i,\,SITE_j} + b_{i,\,(TREE(SITE))_{l(j)}} + b_{i,\,SIZE*TREE(SITE)_{kl(j)}} + b_{i,\,FRUIT_{m(jkl)}}$$

Where

β_1 parameter inversely related to the upper asymptote;

β_2 parameter that relates the upper asymptote (UA) to the lower asymptote (LA), through the function: $e^{UA/LA}$;

β_3 parameter related to the growth rate from the initial values to the final values;

$\beta_{i,\,SIZE_k}$ effect of the jth size for the ith parameter of the fixed effect (i: β_1, β_2, β_3; k: small, medium and big);

$b_{i,\,SITE_j}$ effect of the jth site for the ith parameter of the random effect with $b_{i,\,SITE_j}$ ~ N ([0,0,0], D_{TREE});

$b_{i,\,TREE\,(SITE)l(k)}$ effect of the ith tree within the jth place of the ith parameter of the random effect with $b_{i,\,TREE\,(SITE)l(j)}$ ~ N ([0,0,0], D_{TREE});

$b_{i,\,SIZE*TREE(SITE)kl(j)}$ Interaction between the kth size and the lth tree of the jth site for the random effect parameter with $b_{i,\,SIZE*TREE(SITE)kl(j)}$ ~ N ([0,0,0], $D_{SIZE*TREE}$);

$b_{i,\,FRUTO_{m(jlk)}}$ Effect of the mth fruit of the jth site, the kth size and the lth tree of the random effect with $b_{i,\,FRUTO_{m(jlk)}}$ ~ N ([0,0,0], D_{FRUIT});

$\mathcal{E}_{jlk}$ Random error with $\mathcal{E}_{jlk}$ ~ N (0, 6^2) y $\overset{\varepsilon}{6}{}^2 = 6^2\, I_{nixni}$ where I_{nixni} is an identity matrix and n_i is the number of repeated measurements per fruit

Alternative models are shown in table 1. The ICI values suggested model 2 as the best option. Then, in order to consider the possible correlation, different models were proposed based on Model 2 (table 2). New ICI values suggested model 2b as the best model.

Model	Fixed effects	Random effects	AIC	BIC	logLik
1	Size	Fruit	53.934	54.056	-2 6.951
2	Size	Fruit, Site	51.986	52.154	-25.971
3	Size	Fruit, Site, Site*Size	60.423	60.682	-30.178
4	Size	Fruit, Tree, Site	54.364	54.577	-27.154

Table 1. Fixed effects, random effects and ICIs obtained for the fitted models.

Model	Correlation	AIC	BIC	logLik
2a	first-order autocorrelation in the errors	50.106	50.281	-25.030
2b	first-order continuous autocorrelation in the errors with DAFB as the time covariate	49.083	49.259	-24.518
2c	first-order continuous autocorrelation in the errors with ln of DAFB as the time covariate	49.825	50.001	-24.889

Table2. Specifications of models 2a, 2b and 2c, and their ICIs.

Model	Model specifications		AIC	BIC	logLik
2_1	Fixed effects: β_{1TAMAk}, β_{3TAMAk}	Random effects: at Site level $b_{3,SITIOj}$ and Fruit level $b_{1,FRUTOm(j)}$, $b_{2,FRUTOm(j)}$, $b_{3,FRUTOm(j)}$	92829	92989	-46394
2_2	Idem model 2_1 but considering different variances for each TYPE in the error term		92238	92382	-46101
2_3	Idem Model 2_2, but includes a first-order continuous autocorrelation structure on log(DAFB) in the error term		90584	90735	-45273
2_4	fixed effect: β_{1TAMAk}, β_{3TAMAk}	random effect: at SITE level $b_{TRANS,3,SITEj}$, $b_{LONG,3,SITEj}$ and at FRUIT level $b_{LONG,1,FRUITm(j)}$, $b_{LONG,2,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$	92559	92719	-46259
2_5	fixed effects: β_{1SIZEk}, β_{3TYPEr}	random effect at SITE level $b_{TRANS,3,SITEj}$, $b_{LONG,3,SITEj}$ and at FRUIT level $b_{LONG,1,FRUITm(j)}$, $b_{LONG,2,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$	92612	92685	-46262
2_6	Idem Model 2_5 but considering different variances for each TYPE in the error term.		93974	94085	-46973

Table 3. Specifications of models 2_1, 2_2, 2_3, 2_4, 2_5 and 2_6 and their ICIs.

Transversal data entry

The new candidate models incorporated an additional category in SIZE and a variable called TYPE, discriminating between longitudinal and transversal data (table 3). Using both transversal and longitudinal measurements aimed to leverage strengths of each data source. On one hand, longitudinal measurements contribute to estimating shape parameters. On the other hand, transversal measurements improve variability estimation of random effects, generalizing inference to a broader population.

The best model was Model 2_3, even with poor predictive performance while Model 2_4 presented low significance in some terms. Therefore, Model 2_5 and Model 2_6 were the best candidates. However, inconsistencies between the PC and IC have already been found in other studies modeling error covariance-variance structure (Huang and Meng, 2009; Meng and Huang, 2012). These researches emphasize the purpose of judging the candidate models, *i.e.* if predictive models are sought, using CI is more convenient, while if predictions are sought, PC constitutes a better option

In this study, the random effect of the fruit competed with modelling the residual structure of the error when considering serial correlation over time for the same fruit, generating the inconveniences obtained between ICIs and PCs. A similar situation occurred when estimating more than one random effect from fruits singly measured (transversal data). For this, we proposed a modelling strategy generating differential random effects (b_i) considering data nature as in Models 2_4, 2_5 and 2_6 (table 4). The best model with the best parsimony was Model 2_10, followed by Model 2_11.

Model	Fixed effects	Random effects	AIC	BIC	logLik
2_7	β_{1TYPEr}, β_{3TYPEr}	Site: $b_{TRANS,3,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$	167340	167428	-83659
2_8	β_{1TYPEr}, β_{3TYPEr}	Site: $b_{LONG,3,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,31,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$	179950	180030	-89965
2_9	β_{1TYPEr}, β_{3TYPEr}	Site: $b_{TRANS,3,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$	171105	171193	-85541
2_10	β_{1TYPEr}, β_{3TYPEr}	Site: $b_{LONG,3,SITEj}$, $b_{TRANS,3,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$	108509	108604	-54242
2_11	β_{1TYPEr}, β_{3TYPEr}	Site: $b_{2,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$	116460	116548	-58219

Table 4. Specifications of models 2_7, 2_8, 2_9, 2_10 and 2_11 and their ICIs.

Environmental covariates

The incorporation of environmental covariates explaining b_i at site level was evaluated to increase model predictive capacity. The TCIs were considered at two moments of fruit growth, after full bloom and during the month of December (65 to 95 DAFB). Previous studies have shown the differential effect generated by TCIs, depending on the phenological moment of fruit growth. High temperatures in the early stages of growth produce larger fruits, whereas high temperatures in December result in smaller fruits (Tassile *et al*., 2005).

Considering all TCIs, a variable selection process was used based on their relationship with the empirical Bayesian estimators of $b_{LONG,3,SITEj}$ and $b_{TRANS,3,SITEj}$ of model 2_10 and model 2_11. A total of three TCIs were selected: i) daily temperature accumulation exceeding 16 degrees from 46 DAFB to 100 DAFB (COV_1) in early stages of fruit growth; ii) daily temperature accumulation exceeding 22 degrees from 122 DAFB to 150 DAFB (COV_2) in December; iii) daily temperature accumulation exceeding 17 degrees from 60 DAFB to 100 DAFB (COV_3) in early stages of fruit growth.

Many models were proposed (table 5), all of them having β_{1TYPEr}, β_{3TYPEr} as fixed effects. The selected model parameters are shown on table 6 and figure 1. The models that include TCIs require information after the common moment of diameter measurement. Measurements are made from 65 DAFB to 79 DAFB, while some TCI contemplated in the proposed models can only be obtained from 95 DAFB onwards. For this reason, forecasting between 80 and 95 days after full bloom, the models with TCI cannot be used, while if the forecast is made after 95 DAFB, models including them can be used. However, we observed that model predictive capacity not including TCIs depends to a great extent on the prediction moment, while models including TCI showed high predictive performances in all situations considered.

Model	Fixed effects	Random effects
2_12	β_{1TYPEr}, β_{3TYPEr} COV_1 for β_3	Site: $b_{LONG,3,SITEj}$, $b_{TRANS,3,SITEj}$ Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j}$
2_13	β_{1TYPEr}, β_{3TYPEr} COV_1 for β_3	Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,31,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_14	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3	Site: $b_{LONG,3,SITEj}$ $b_{TRANS,3,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_15	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3	Site: $b_{2,SITEj,}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_16	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3, COV_3 for β_2	Site: $b_{LONG,3,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_17	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3, COV_3 for β_2	Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_18	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3, COV_3 for β_2	Site: $b_{LONG,3,SITEj}$, $b_{TRANS,3,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_19	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3, COV_3 for β_2	Site: $b_{LONG,3,SITEj}$, $b_{TRANS,3,SITEj}$, , $b_{2,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_20	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3, COV_3 for β_2	Site: $b_{2,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$
2_21	β_{1TYPEr}, β_{3TYPEr} COV_2 for β_3, COV_3 for β_2	Site: $b_{LONG,3,SITEj}$, $b_{2,SITEj}$; Fruit: $b_{2,FRUITm(j)}$, $b_{LONG,1,FRUITm(j)}$, $b_{LONG,3,FRUITm(j)}$, $b_{TRANS,3,FRUITm(j)}$

The incorporation of environmental covariates is of the form: βi+COV_i*TCI.

Table 5. Specifications of models 2_12, 2_13, 2_14, 2_15, 2_16, 2_17, 2_18, 2_19, 2_20 and 2_21.

	Parameters			Model estimations			
	Model 2_12	Model 2_14	Model 2_15	Model 2_16	Model 2_17	Model 2_20	Model 2_21
$\beta_{1INTERCEPT}$	-5.501**	-5.508**	-5.858**	-5.575**	-5.553**	-5.534**	-5.558**
β_{1LONG}	6.671**	6.678**	7.030**	6.748**	6.729**	0.565**	6.731**
$\beta_{2INTERCEPT}$	8.752**	8.757**	8.875**	9.305**	9.372**	9.401**	9.418**
$\beta_{3INTERCEPT}$	10.001**	9.982**	9.984**	9.982**	9.982**	9.989**	9.984**
β_{3LONG}	-0.242**	-0.245**	-0.248**	-0.246**	-0.246**	-0.247**	-0.246**
α_{2TEMP}				-0.00168 (3)**	-0.01729(3)**	-0.01845(3)NS	-0.01934NS
α_{3TEMP}	-0.00044(1)**	-0.00004(2)NS	-0.00058(2)NS	-0.00001(2)NS	-0.00006(2)NS	-0.00616(2)**	-0.00018(2)NS
$\sigma_{2,SITIOj}$	0.033	0.002	8.263	.	.	0.531	0.524
$\sigma_{bTRANS,3,SITEj}$	0.033	0.002	.	0.040	.	.	.
$\sigma_{bLONG,3,SITEj}$	0.021	0.495	.	.	0.041	.	0.001
$\sigma_{bLONG,1,FRUITm(j)}$	0.051	0.035	0.042	0.038	2.954	0.039	0.038
$\sigma_{b,2,FRUITm(j)}$	1.672	1.805	1.646	2.113	.	2.052	2.087
$\sigma_{bLONG,3,FRUITm(j)}$	0.001	0.045	0.297	0.001	0.002	0.001	0.001
$\sigma_{bTRANS,3,FRUITm(j)}$	0.079	0.007	0.021	0.004	0.003	0.004	0.033
σ_{RES}	0,545	0,554	0,529	0.549	0.533	0.535	0.525

$\sigma_{bX,1i}$, $\sigma_{bX,2,i}$ and $\sigma_{bX,3,i}$ are random effects at SITE level. $\sigma_{bX,1ij}$, $\sigma_{bX,2,ij}$ and $\sigma_{bX,3,ij}$ are random effects at FRUIT level.
(1) COV_1; (2)COV_2; (3)COV_3; (NS)Not significance; (**)Significance ($p<0.01$). Test based on Wald statistics with normal asymptotic distribution.

Table 6. Parameter estimation for fitted models.

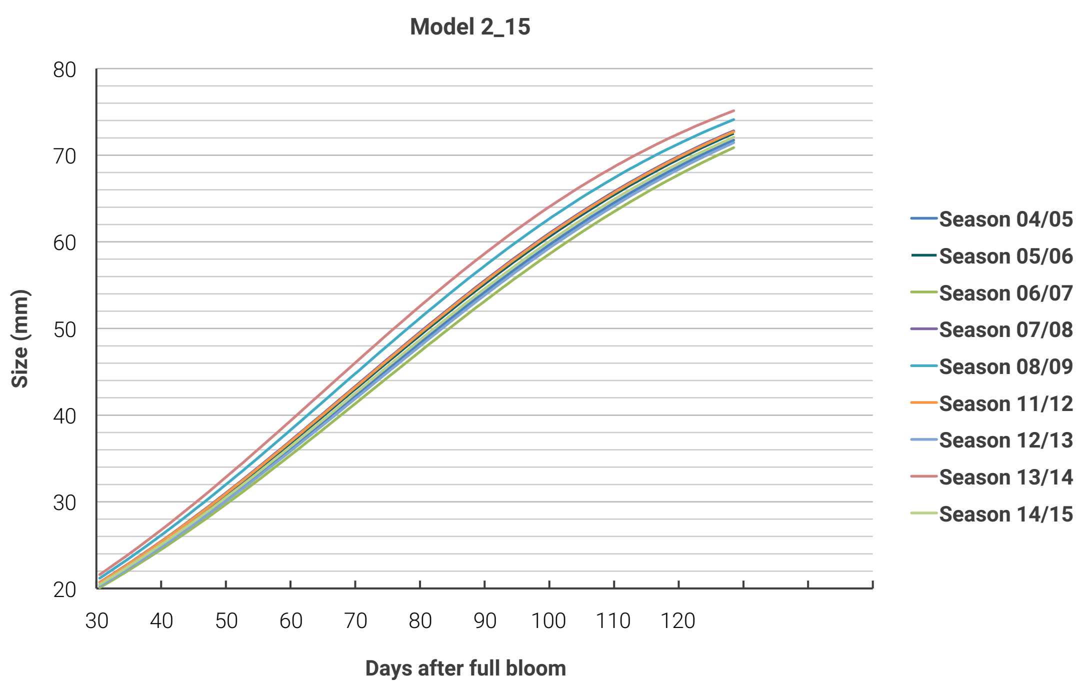

Figure 1. Fruit growth model 2_15 for seasons in the modeling base for longitudinal data. Shows the fruit sizes (mm) as a function of days after full bloom.

The good performance of models contemplating TCI can be explained by temperature effects on cell expansion throughout December and, therefore, in final fruit size obtained, in accordance with Tassile *et al*. (2005) and Giménez, (2012).

Generally, variables expressing temporal or local spatial fluctuations increase accuracy and reduce NLME prediction bias. The results obtained in this work agreed with those found in other study areas such as forester (Crecente Campo *et* *al*., 2010; Calama and Montero, 2005; Newton and Amposah, 2007); in pharmacodynamics (Mandema *et al*., 1992; Wakefield, 1996; Wahlby, 2002; Ribbing and Jonsson, 2004; Bonate, 2006; Bertrand *et al*., 2011; Hagenbunch, 2011), in epidemiology (Distiller *et al*., 2010); in animal production (Furtado Campos *et al*., 2014). Also, previous results on the differential effect of thermal accumulation on fruit phenological moments were corroborated. This is how in early growth stages, high temperatures stimulate cell growth and fruit potential size, while at the end of the cycle high temperatures negatively affect growth rates (Tassile *et al*., 2005; Tassile *et al*., 2006; Rodriguez, 2009).

CONCLUSION

Multilevel NLMM showed the advantage of discerning and quantifying stochastic variability sources at different levels, obtaining better ICI and CP in comparison to models considering a single level of variability in random effects.

The candidate models resulting from modelling both transversal and longitudinal data were multilevel models with random effects at site and fruit level. In addition, the capacity of some multilevel models without considering TCIs to achieve high ICI was very interesting. This occurred due to the ability of site-level random effects to account for stochastic variability without requiring measurement or identification of the variation source.

The incorporation of TCIs characterizing different moments of fruit growth to the NLMM was favorable.

REFERENCES

ALVAREZ, O.; BOCHE, S. 1999. Modelos matemáticos para describir crecimientos doble sigmoideos en frutos de Nectarín tardío (cv. Sun Grand). Agro sur, 27:1, 21-27.

ATAY, E.; PIRLAK, L.; ATAY, A. N. 2010. Determination of fruit growth in some apple varieties. Journal of Agricultural Sciences, 10, 1-8. https://doi.org/10.1501/Tarimbil_0000001115

AVANZA, M.M. 2010. Desarrollo de una propuesta metodológica para la predicción de cosecha de naranjo dulce (*Citrus sinensis L. Osbeck*) var Valencia late en la provincia de Corrientes, Argentina. Tesis para acceder al título de Doctor en el Área de Recursos Naturales, Universidad Nacional del Nordeste. 230 p.

BERTRAND, J.; LAFFONT, C.M.; MENTRE, F.; CHENEL, M.; COMETS, M. 2011. Development of a complex parent-metabolite joint population pharmacokinetic model. The American Association of Pharmaceutical Scientists Journal, 13:3, 390-404. https://doi.org/10.1208/s12248-011-9282-9

BONATE, P.L. 2006. Pharmacokinetics pharmacodynamic modeling and simulation. Springer, New York.

BRAMARDI, S.J.; CASTRO, H.R.; ZANELLI, M.L. 1997 Fruit growth pattern of pear cv. Bartlett and Packham´s Triumph to improve hand thinning. Acta Horticulturae, 475, 283-293.

CALAMA, R; MONTERO, G. 2005. Multilevel linear mixed model for tree diameter increment in stone pine (Pinus pinea). Silva Fenn, 39, 37-54.

CARRERO, O.; JEREZ, M.; MACCHIAVELLI, R.; GIAMPAOLO, O.; STOCK, J. 2008. Ajuste de curvas de índice de sitio mediante modelos mixtos para plantaciones de Eucalyptus urophylla en Venezuela. Interciencia, 33:4, 265-272.

CRECENTE CAMPO, F.; TOME, M.; SOAERES, P.; DIEGUEZ ARAND, U. 2010. A generalized nonlinear mixed-effects height-diameter model for Eucalyptus globulus l. in northwestern Spain. Forest Ecology and Management, 259, 494-952.

DISTILLER, G.; LITTLE, F.; BARNES, K. 2010. Nonlinear mixed effects modeling of gametocyte carriage in patients with uncomplicated malaria. Malaria Journal, 60, 1-12.

FURTADO CAMPOS, H.; NOBLET, J.; JAGUELIN-PEYRAUD, Y.; GILBERT, H.; MORMEDE, P.; MIRANDA, F.; RENAUDEAU, D. 2014. Thermoregulatory responses during thermal acclimation in pigs divergently selected for residual feed intake. International Journal of Biometeorology, 8, 1545-1557.

GALAN, C.; VAZQUEZ, L.; GARCIA-MOZO, H.; DOMINGUEZ, E. 2004. Forecasting olive (*Olea europaea*) crop yield based on pollen emission. Field Crops Research, 86, 43-51. https://doi.org/10.1016/S0378-4290(03)00170-9

GIMENEZ, G.N. 2012. Determinación de los patrones de crecimiento y abscisión de frutos para ajustar el momento oportuno de raleo en perales cv. 'Williams'. Tesis para acceder al título de Magister en Fruticultura de Clima Templado-Frío. Universidad de Bologna y Universidad Nacional del Comahue. 128 p.

GIMENEZ, G.N.; DUSSI, M.C.; ZON, K.; NIEKI, J.; SZABO, Z.; RACSKO, J.; REEB, P.D. 2010. Fruit growth and abscission pattern of 'Williams' pear treated with benzyladenine. Acta Horticulturae, 884, 481-490. https://doi.org/10.17660/ActaHortic.2010.884.61

GLENN, D.M.; TABB, A. 2019. Evaluation of Five Methods to Measure Normalized Difference Vegetation Index (NDVI) in Apple and Citrus. International Journal of Fruit Science, 19:2, 191-210. https://doi.org/10.1080/15538362.2018.1502720

GODOY, C.; MONTERUBBIANESI, G.; TOGNETTI, J. 2008. Analysis of highbush blueberry (Vaccinium corymbosum L.) fruit growth with exponential mixed models. Scientia Horticulturae, 115, 368-376.

GOUDRIAAN, J.; MONTEITH, L. A. 1990. mathematical function for crop growth based on light interseption and leaf area expansion. Annals of Botany, 66:6, 695-701.

HAGENBUCH, N.A. 2011. Comparison of Four Methods to Analyse a Non-Linear Mixed-Effects Model Using Simulated Pharmacokinetic Data. Tesis de grado, Swiss Federal Institute of Technology Zurich.

HUANG, S.; MENG, S.X. 2009. Estimating a Nonlinear Mixed Volume-age Model with and without Taking into Account Serrially-correlated Errors: Differences and Implications. Modern Applied Science, 3:5, 3-20.

HUANG, S.; MENG, S. 2012 Accounting for serial correlation and its impact on forecasting ability of a fixed and mixed effects basal area model: a case study. European Forest Research, (131):541-552.

JIMENEZ, C.M.; ROYO DIAZ, J.B.R. 2003. Statistical Model Estimates Potential Yields in Pear Cultivars 'Blanquilla' and 'Conference' before Bloom. American Society for Horticultural Science, 128:4, 452-457. https://doi.org/10.21273/JASHS.128.4.0452

KAACK, K.; LINDHARD PEDERSEN, H. 2010. Prediction of Diameter, Weight and Quality of Apple Fruit (*Malus domestica Borkh*.) cv. 'Elstar' using Climatic Variables and their Interactions. European Journal of Horticultural Science, 75:2, 60-70.

KUMAR, A.; SHARMA, P.; AMBRAMMAL, S. K. 2014. Climatic Effects on Food Grain Productivity in India: A Crop Wise Analysis. Journal of Studies in Dynamics and Change, 1:1, 38-48.

LAKSO, A.N.; CORELLI GRAPPADELLI, L.; BARNARD, J.; GOFFINET, M.C.; AN, M.C.G. 1995. An expo linear model of the growth pattern of the apple fruit. Journal of Horticultural Science, 70:3, 389-394. https://doi.org/10.1080/14620316.1995.11515308

LINDSTROM, M.J.; BATES, D.M. 1990. Nonlinear Mixed Effects Models for Repeated Measures Data. Biometrics, 46:3, 673-687. https://doi.org/10.2307/2532087

MAGEIN, H. 1989. Growth and abscission dynamics of 'Cox' s Orange Pippin' and 'Golden Delicious' apple fruits. Journal of Horticultural Science, 64:3, 265-273. https://doi.org/10.1080/14620316.1989.11515954

MANDEMA, J.; VEROTTA, D.; SCHEINER, L. 1992. Building population pharmacokinetics/ pharmacodymanic models. Models for covariate effects. Journal of Pharmacokinetics and Biopharmaceutics, 20, 511-529. https://doi.org/ 10.1007/BF01061469

MCARTNEY, S.; PALMER, J.W.; ADAMS, H.M. 1996. Crop loading studies with 'Royal Gala' and 'Braeburn' apples: Effect of time and level of hand thinning. New Zealand Journal of Crop and Horticultural Science, 24:4, 401-407. https://doi.org/10.1080/01140671.1996.9513977

MENG, S.; HUANG, S. 2012. Accounting for serial correlation and its impact on forecasting ability of a fixed and mixed effects basal area model: a case study. European Forest Research,131:3, 541-552. https://doi.org/:10.1007/s10342-011-0527-z

NEWTON, P.F.; AMPOSAH, I.G. 2007. Comparative evaluation of five height-diameter models developed for black spruce and jack pine stand-types in terms of goodness-of-fit, lack-of-fitand predictive ability. Forest Ecology and Management, 247:1, 149-166. https://doi.org/10.1016/j.foreco.2007.04.029

PASQUA PAULINO, S.E.; ALVES MOURÃO FILHO, F.; NUNES MAIA, A.; CANTUARIAS AVILÉS, T.E.; NETO, D.D. 2007. Agrometeorological models for 'Valencia' and 'Hamlin' sweet oranges to estimate the number of fruits per plant. Science Agriculture, 64:1, 1-11. https://doi.org/10.1590/S0103-90162007000100001

PINHEIRO, J.C.; BATES, D.M. 2000. Mixed-Effects Models in S and S-Plus. Springer, New York.

RATKOWSKY, D.A. 1993. Principles of nonlinear regression modeling. Journal of Industrial Microbiology, 12, 195-199. https://doi.org/10.1007/BF01584190

R CORE TEAM R. 2018. A Language and Environment for Statistical Computing, R Foundation for Statistical Computing. (Available at: http://www.R-project.org verified on July 17, 2023).

RIBBING, J.; JONSSON, E.N. 2004. Power, selection bias and predictive performance of the population pharmacokinetic covariate model. Journal of Pharmacokinetics and Pharmacodynamics, 31, 109-134.

RODRIGUEZ, A.B. 2009. Influencia de la Temperatura del Aire y del Raleo sobre el Tamaño del Fruto de Peral (Pyrus communis) cv Williams durante sus Distintas Fases de Desarrollo. Tesis para acceder al título de Magíster en Ciencias Agropecuarias, Universidad Nacional de Córdoba. 195 p.

TASSILE, V. 2020. Aplicación de modelos mixtos no lineales en el crecimiento de frutos de peral (*Pyrus communis*) cv. Williams. Tesis para acceder al título de Magister en Estadística Aplicada, Universidad Nacional de Córdoba. 220 p.

TASSILE, V.; ALVAREZ, O.; BRAMARDI, S. 2002. Aplicación de modelos nolineales mixtos en la descripción del crecimiento de frutos de pepita. Posters section at the V Congreso Latinoamericano de Sociedades de Estadística. Buenos Aires, Argentina. 133 p.

TASSILE, V.; BRAMARDI, S.J.; MACCHIAVELLI, R. 2015. Modelos mixtos obtenidos a partir de información longitudinal y transversal. Posters section at the I Congreso Argentino de Estadística. Universidad Nacional de Córdoba, Córdoba.

TASSILE, V.; BRAMARDI, S.J.; REEB, P.D.; DE BERNARDIN, F. 2005. Modelos de crecimiento de frutos de pepita: mejora en la precisión de las estimaciones a través de la inclusión de covariables ambientales y componentes aleatorias. Posters section at the X Reunión Científica del Grupo Argentino de Biometría. Universidad Nacional del Nordeste, Corrientes.

TASSILE, V.; BRAMARDI, S.J.; REEB, P.D.; DE BERNARDIN, F. 2006. Validación de modelos no lineales mixtos y su aplicación con fines de pronóstico. Posters section at the VII CLATSE, Rosario.

WAHLBY, U. 2002. Methodological Studies on Covariate Model Building in Population Pharmacokinetic/Pharmacodynamics Analysis. Tesis para acceder al título de Doctor, Uppsala University.

WAKEFIELD, J. 1996. The bayesian analysis of population pharmacokinetic models. Journal of the American Statistical Association, 91, 62-75. https://doi.org/10.2307/2291383

YANG, Y.; HUANG, S. 2011. Comparison of different methods for fitting nonlinear mixed forest models and for making predictions. Canadian Journal of Forest Research, 41:8, 1671-1686.

https://doi.org/10.1139/X11-071

ZADRAVEC, P.; VEBERIC, R.; STAMPAR, F.; SCHMITZER, V.; ELER, K. 2014. Fruit Growth Patterns of Four Apple Cultivars Using Nonlinear Growth Models. European Journal Horticulturae Science, 79:2, 52-59.

ORTEGA-FARIAS, S.; FLORES, L.; RETAMALES, S. 1998. Modelo logístico para predecir el crecimiento en diámetro de las manzanas (variedad Granny Smith). Revista Frutícola, 19, 15-18.